DOCUMENT RESUME

ED 156 534 SE 024 573

TITLE Graphs and Statistics: A Resource Handbook.
INSTITUTION New York State Education Dept., Albany. Bureau of General Education Curriculum Development.; State Univ. of New York, Albany.
PUB DATE 77
NOTE 90p.; Contains occasional small print

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS Activity Units; Data Collection; Data Processing; *Graphs; Instructional Materials; *Resource Guides; Secondary Education; *Secondary School Mathematics; Statistical Data; *Statistics; *Tables (Data)

ABSTRACT

Graphical representation of statistical data is the focus of this resource handbook. Only graphs which present numerical information are discussed. Activities involving the making, interpreting, and use of various types of graphs and tables are included. Sections are also included which discuss statistical terms, normal distribution and probability, and collecting and preparing statistical data. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED156534

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Joan Miller

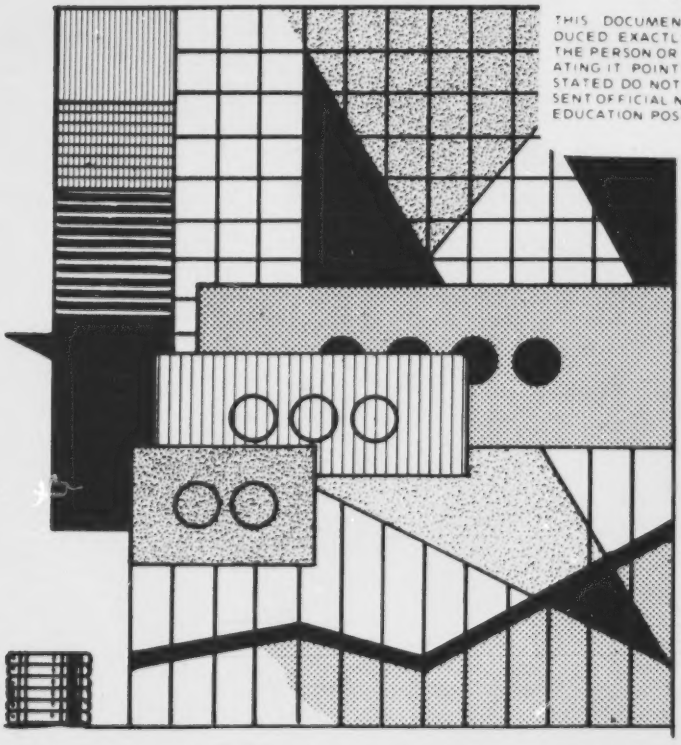
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

GRAPHS and STATISTICS

A Resource Handbook

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY



The University of the State of New York
THE STATE EDUCATION DEPARTMENT
Bureau of General Education Curriculum Development
Albany, New York

2

SE 024 573

GRAPHS AND STATISTICS

A RESOURCE HANDBOOK

(Reprint November 1977)

The University of The State of New York
THE STATE EDUCATION DEPARTMENT
Bureau of General Education Curriculum Development
Albany, New York 12234

THE UNIVERSITY OF THE STATE OF NEW YORK

Regents of The University (with years when terms expire)

1981 Theodore M. Black, A.B., Litt.D., LL.D., Pd.D., D.C.L., L.H.D., Chancellor - - - - - - - - - - - - - - - - - - - Sands Point
1987 Carl H. Pforzheimer, Jr., A.B., M.B.A., D.C.S., H.H.D., Vice Chancellor- Purchase
1978 Alexander J. Allan, Jr., LL.D., Litt.D. - - - - - - - - - Troy
1986 Kenneth B. Clark, A.B., M.S., Ph.D., LL.D., L.H.D., D.Sc.- Hastings on Hudson
1983 Harold E. Newcomb, B.A. - - - - - - - - - - - - - - - - Owego
1988 Willard A. Genrich, LL.B., L.H.D., LL.D.- - - - - - - - Buffalo
1982 Emlyn I. Griffith, A.B., J.D. - - - - - - - - - - - - - Rome
1983 Mary Alice Kendall, B.S.- - - - - - - - - - - - - - - - Irondequoit
1984 Jorge L. Batista, B.A., J.D., LL.D. - - - - - - - - - - Bronx
1982 Louis E. Yavner, LL.B.- - - - - - - - - - - - - - - - - New York
1979 Laura B. Chodos, B.A., M.A. - - - - - - - - - - - - - - Clifton Park
1980 Martin C. Barell, B.A., I.A., LL.B. - - - - - - - - - - Kings Point
1981 Joseph R. Bongiorno, B.S., M.D. - - - - - - - - - - - - Brooklyn
1984 Louise P. Matteoni, B.A., M.A., Ph.D. - - - - - - - - - Bayside
1981 J. Edward Meyer, B.A., LL.B.- - - - - - - - - - - - - - Chappaqua

President of The University and Commissioner of Education
Gordon M. Ambach

Executive Deputy Commissioner of Education

Acting Deputy Commissioner for Elementary, Secondary and Continuing Education
John W. Polley

Associate Commissioner for Instructional Services
Vivienne Anderson

Assistant Commissioner for General Education and Curricular Services
Maria Ramirez

Director, Division for Curriculum Services
Gordon E. Van Hooft

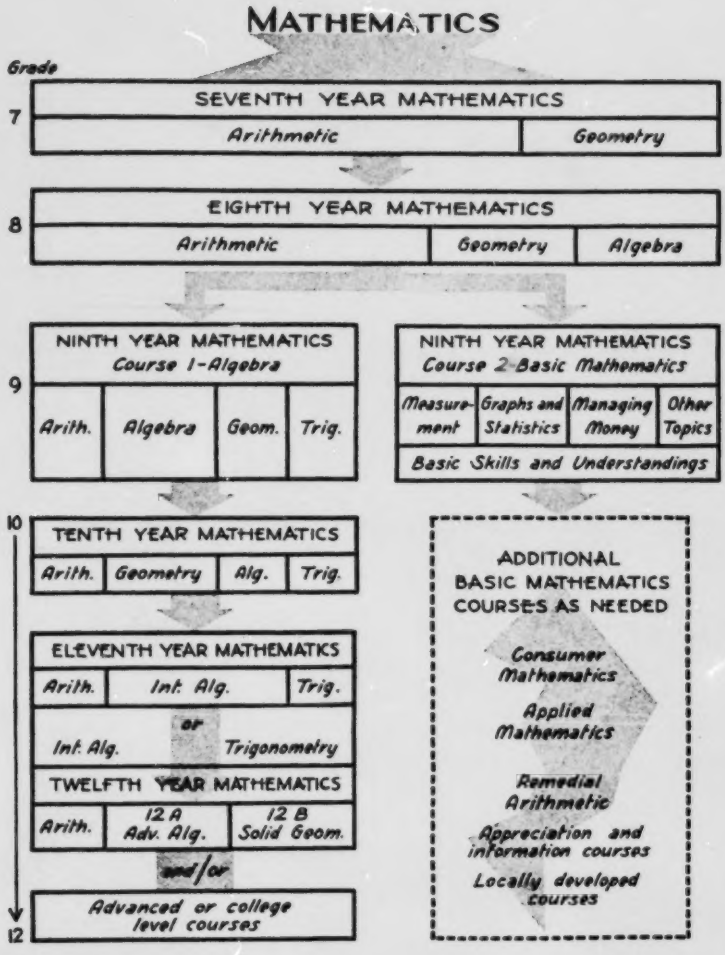
Chief, Bureau of General Education Curriculum Development
Herbert Bothamley

Director, Division of General Education
Ted T. Grenda

Chief, Bureau of Mathematics Education
Fredric Paul

CONTENTS

		Page
Introductory Activities		1
I.	Pictographs and Bar Graphs	5
II.	Line Graphs	17
III.	Circle Graphs and Composite Graphs	36
IV.	Scatter Diagrams	50
V.	Statistical Tables	55
VI.	Statistical Terms	60
VII.	Normal Distribution and Probability	74
VIII.	Collecting and Preparing Statistical Data	78


MATHEMATICS
Grade
7
SEVENTH YEAR MATHEMATICS
Arithmetic
Geometry
8
EIGHTH YEAR MATHEMATICS
Arithmetic
Geometry
Algebra
9
NINTH YEAR MATHEMATICS
Course 1-Algebra
Arith.
Algebra
Geom.
Trig.
NINTH YEAR MATHEMATICS
Course 2-Basic Mathematics
Measurement
Graphs and Statistics
Managing Money
Other Topics
Basic Skills and Understandings
10
TENTH YEAR MATHEMATICS
Arith.
Geometry
Alg.
Trig.
ELEVENTH YEAR MATHEMATICS
Arith.
Inf. Alg.
Trig.
Inf. Alg.
or
Trigonometry
TWELFTH YEAR MATHEMATICS
Arith.
12 A Adv. Alg.
12 B Solid Geom.
and/or
12
Advanced or college level courses
ADDITIONAL BASIC MATHEMATICS COURSES AS NEEDED
Consumer Mathematics
Applied Mathematics
Remedial Arithmetic
Appreciation and information courses
Locally developed courses

Introductory Activities

Limitation of Scope

When information is presented graphically it is more readily assimilated. The number of graphs which appear in newspapers, periodicals and textbooks indicates the usefulness of graphs as an instructional medium. In a broad sense, anything which "pictures" information could be considered a graphical representation. For example, a photograph is a representation of the intensity of reflected or direct light on a surface. Similarly a roadmap is a graphical representation of information. Within this pamphlet, however, graphs will be considered in a more limited scope. Only graphs which present numerical information will be discussed.

1. Building a Collection of Illustrative Graphs

Since simple illustrations of the various types of graphs are difficult to locate when the need for them arises in the classroom, it is essential that each teacher develop his own collection. Throughout the year graphs in magazines, newspapers and other sources available to the teacher should be examined critically and those which are considered suitable added to the collection. Each such graph ought then to be keyed as to type and purpose and filed for future use. Periodically the collection should be reviewed in order to eliminate obsolete material and to prevent duplication.

Variation: Since some published graphs violate many of the principles of good graphing, it is worthwhile to build a collection of poor graphs. For instance, bar graphs which do not begin at zero or which fail to have an accurate title are far too plentiful.

2. Reproducing Materials for Interpretation

Sometimes it is desirable that each pupil have a duplicate copy of graphs or tables clipped from newspapers, magazines or commercial or industrial sources. Copies can be made on a photographic copier, which is sometimes available in school business offices. Copies will be the same size as the original, but are made on sheets 8½" x 11" or 8½" x 14"; to save space, two or more small items may be copied on the same sheet.

These copies may be kept and used with different classes until such time as the data become out-of-date and the copies should be discarded.

The cost of each sheet is approximately 2½ cents. This need not be considered prohibitive if it is equated with the time and effort needed to make a master copy for duplicating.

3. Displaying Materials for Interpretation

In order to have the whole class join in discussion of graphs or tables, use an opaque projector to throw the picture on a plain, light-colored surface in the room. Then the important points of the graph can be indicated with a ruler, either by the teacher or by a pupil. This is useful for displaying graphs collected from magazines or newspapers if they are too small for easy visibility in the average classroom.

If an opaque projector is not available, make several demonstration graphs on sheets of paper, 18" x 24", illustrating the different types of graphs to be interpreted.

Materials which are appropriate and which can be bought at an art supply store or stationers include:

Pads of newsprint, 18" x 24"
Artists' sketch pads
Illustration or poster board
Large sheets of graph paper

Use a marking pen in various colors for better visibility of the graph from the back of the classroom.

Once these graphs have been prepared, they may be kept from year to year, with a few substitutions to keep the data current. Graphs on lightweight paper may be taped to old window shades and rolled up for storage. If brackets are available over the blackboard, the shade can be attached there and pulled down when the graph is needed.

Graphs constructed on stiff paper may be displayed on an easel, a metal stand or the chalk-tray. They may have holes punched in three places, and be hung, either for storage or display, on hooks such as are used for cups in a cupboard; or they may be clipped to a wire across the top of the blackboard.

4. Stimulating Class Understanding of Graphs or Tables

A planned order of procedure gives more direction to class understanding of a graph or table than using haphazard questions.

5. Aiding Students in Interpreting Graphs

It is useful to have a duplicated form on which pupils may record their readings when they are first learning to interpret graphs. Such a sheet is shown on page 9. This can be typed double-spaced on a sheet of paper, 8½" x 11".

Each graph assigned may be given an identification number as well as a title. References included in the following form are to the graph shown for activity 21.

Name _____

Graph Number: _____	Graph Number: _____
Item: Automobiles Caption	Percent of Families in U. S. Buying Explanation
Item: _____ Caption	_____ Explanation
Data: 1946—11% 1947—15% 1948—17% 1949—22% 1950—24% 1951—22% 1952—21%	Data:
Relationship or Trend: Increasing with peak in 1950, then decreasing.	Relationship or Trend:
Related Factors: (1) Cars are usually traded in every few years. (2) During World War II, cars were not produced; postwar models were not sold in volume until 1946.	Related Factors:
Comparisons: All items shown increased until about 1950, then dropped. Many more families buy cars than anything else shown on the graph, in any year.	Comparisons:

The following steps, in order of difficulty, may serve as a guide:

1. Identifying items pictured or tabulated through use of title, subtitle, captions and key
2. Understanding scales used and noting any irregularities
3. Reading graph or table for any single item shown
4. Comparing any two items shown
   a. Amount more or less (addition or subtraction)
   b. Number of times as great as or fractional part one is of another (multiplication or division)

5. Combining or grouping items to obtain information not given specifically
6. Recognizing and stating trends or making a generalization indicated by the graph or table
7. Trying to explain notable features of a graph or table by considering related factors not given
8. Having students construct their own graphs using data from any of the sources suggested in section VIII

Caution: In regard to 4*b* above, avoid the expression, "number of times greater than." Although it is correct to say "40 is 3 times greater than 10," it would be clearer to say "40 is 4 times as great as 10." It is not correct to say "40 is 4 times greater than 10."

In like manner, we can say 40 is 300 percent greater than 10, or 40 is 400 percent of 10, but we cannot correctly say that 40 is 400 percent more than 10.

Section I

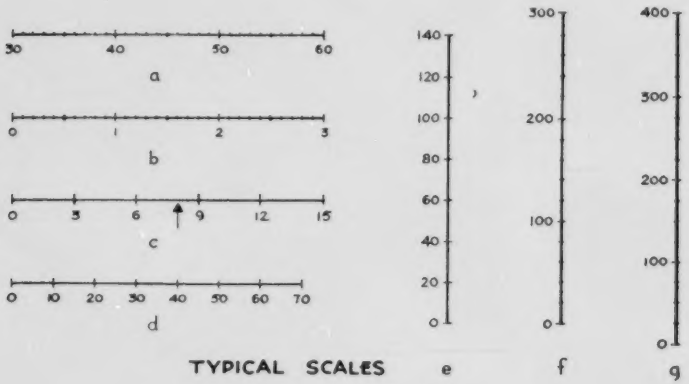
Pictographs and Bar Graphs

6. Making Demonstration Scales

To give practice in estimating the scale values of points between division lines on a scale, prepare several large model scales that may be seen from all parts of the room.

White shelf paper 13 inches wide can be obtained in 100-foot rolls at grocery stores; a strip of this paper 6½ inches wide and 5 or 6 feet long can be inked with a marking pen to show a typical scale. Attach loops of cellophane tape to the back of the strip to display the scale on the chalkboard, bulletin board or wall. Both horizontal and vertical scales may be prepared.

Some typical scales to consider are shown in diagrams *a* through *g* in the order of their complexity.



TYPICAL SCALES

7. Reading Scales

To find the size of subdivisions, find the difference between two consecutive numbers on the scale, and divide this by the number of spaces between the two numbers. For example, in *b* the consecutive numbers differ by 1 and there are 10 spaces between the numbers marked; therefore, each space represents $1 \div 10$, or $\frac{1}{10}$, or .1. Use a pencil point to indicate various readings on the lines, and have pupils give the value correct to the nearest tenth.

To practice reading "between the lines," identify the scale values of the two lines between which the pointer falls, and find the difference between them. Find the fractional part of this distance, which is included between the lower number and the point to be read; compute this fraction of the difference between the numbers and add this to the smaller value.

In *c*, for instance, the arrow falls between 6 and 9, the difference between these is 3. Approximately $\frac{2}{3}$ of this distance is included between the lower number and the arrow; $\frac{2}{3}$ of 3 is 2; $6 + 2 = 8$. The arrow is at 8 on the scale.

Focus on successively harder points and discuss how each should be read. A set of five readings at a time may be asked for; each student will write down all five answers, and one student may be asked to give all five of his readings. Practice may be continued at intervals until the pupils can read a set of five points correctly most of the time.

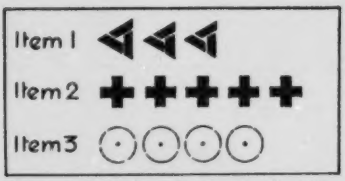
8. Using Pictographs

Description:

Type 1. One or more rows of identical figures, or symbols, and a key which indicates the unit which each figure or symbol represents. (See *a*.)

Type 2. Two or more similar figures or symbols of different size, labeled and identified for comparison, as in *b*.

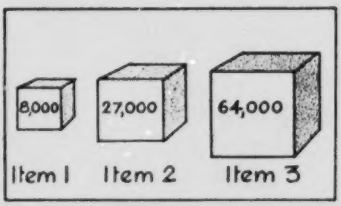
Uses: To show comparative size. To show simple statistics in an appealing manner.



Each symbol represents 1 million

a

b



c

Type 1 is used when several items are to be shown and the ratios of the sizes represented are smaller than 4:1, as in *a*.

Type 2 is used when there are only a few items to be shown and the ratios are 4:1 or greater, as in *b*. A variation of this second type uses volumes to represent magnitudes. This type is usually reserved for situations where the ratios are 8:1 or greater, as in *c*. (For example, a ratio of 125:1 can be shown by figures whose corresponding dimensions are in the ratio of 5:1, since volumes of similar figures are proportional to the cube of the ratio of the corresponding sides.

Notes: Spacing of rows should be uniform.

Double or multiple rows of symbols may be used to reduce the length of a row.

The starting points for all rows should lie in a vertical line.

Different sized figures should be proportional to the amounts they represent.

Title and graph must agree.

Interpretation: The amount represented by each row is determined by counting the number of symbols in each row and multiplying this number by the amount each symbol represents.

Comparisons of amounts represented by the rows of symbols are made by counting the number of symbols in each row and comparing these numbers. This graph is similar to a bar graph.

Comparisons of amounts represented by different sized figures are made by determining whether the figures vary according to one, two or three dimensions and then comparing either lengths, areas, or volumes.

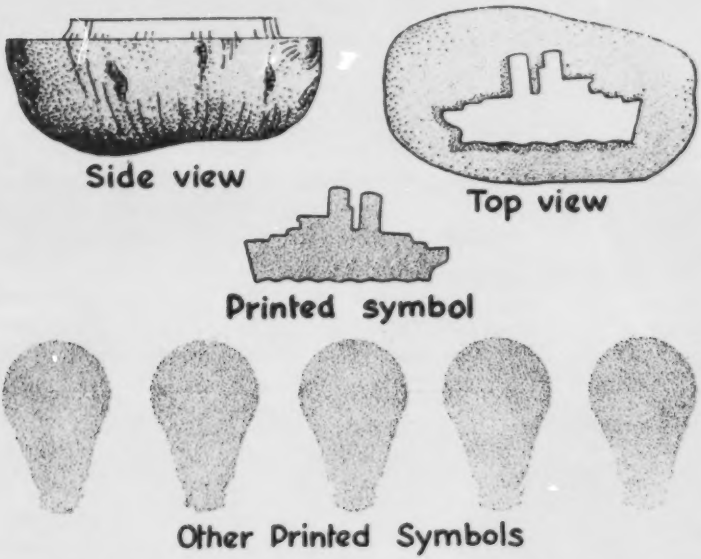
Materials for the Construction of Pictographs. Decorative symbols can be purchased by the hundred at a slight cost in most stationery shops. These may be used as symbols in the construction of the pictographs. Symbols can also be cut from colored construction paper.

Symbols on adhesive-backed sheets can be obtained commercially, at higher cost. An art supply store can provide information on these items.

9. Making a Potato Stamp

A potato can be used to make a stamp for printing large sized figures on a demonstration bar-type pictogram. See diagram on page 14.

First, cut the potato through the middle, taking care that the cut is straight. Using an indelible pencil, draw the desired symbol on the flat surface of the potato. This design should be in reverse. Cut around the outline with a sharp knife, removing all the surface not included in the design, to a depth of about ¼ inch. Any slight cuts made in the surface of the design will print as white lines on the design. This raised design



may be inked by using the back of a fountain pen, or a stamp pad; it can then be used as a stamp to make symbols for a pictograph.

These stamps can be reused within 24 hours, but will shrivel up after two or three days.

Variations:

Use a cookie-cutter to imprint a design on the cut surface.

A rubber stamp set can be obtained in a toy store; these sometimes include figures which could be used in a pictogram.

An art gum eraser can also be cut to form designs; however, the eraser tends to crumble and break unless care is taken in cutting; and cuts made on its surface will not print as lines, unless some of the gum is removed.

10. Using Flannel-Board Pictographs

In order to demonstrate pictographs, cutouts may be made from pieces of felt and used on a flannel-board.

A flannel-board can be made by stretching dark-colored flannel or felt over a flat surface, so that a type of bulletin board is constructed. The stand for the flannel-board should allow it to lean back about 15° from the vertical, so that the cutouts will not fall off. Commercially made flannel-boards and cutouts are available in a limited variety of forms.

Write the title and captions on the blackboard above or near the flannel-board. Use felt cutouts of numbers and symbols to represent the scale.

Start with a pictograph which compares only two items, and which has a unit scale. Have data which shows small differences and simple multiples. The data and the scale used on the pictograph can easily be changed to vary the difficulty of the readings and the relationships illustrated.

If the teacher has time and materials, the title and key may be made from letters cut out of sandpaper, felt or other materials which will adhere to the flannel-board. In that case it is best to choose a topic for which the captions (Yes—No, or Boys—Girls, or Republicans—Democrats) and much of the title can remain substantially the same, for several graphs.

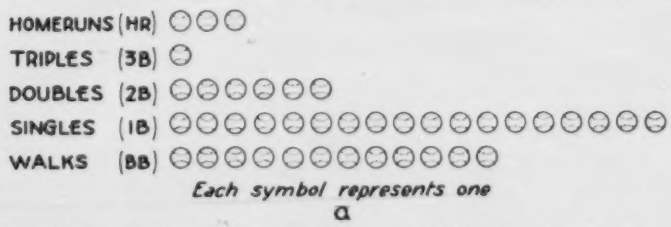
Variation: By using squares or rectangles of felt, bar graphs can also be shown.

11. Interpreting Unit Pictographs and Labeled Bar Graphs

Display graphs *a* and *b* i; classroom.

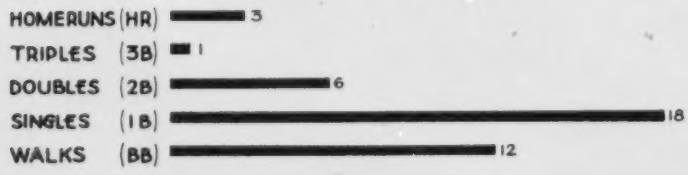
Ask a pupil to read the title, the captions and the keys on both graphs; have the pupil tell what information he would expect to find in each graph. Discuss whether the titles could be shortened in any way.

1956 Baseball Record of Hits and Walks for John Perkins



a

1956 Baseball Record of Hits and Walks for John Perkins



b

15

Call attention to the scale: on a unit pictograph each symbol represents *one*. Have the pupils count or read the magnitude of each item shown on the pictograph in any order. Point out that the same data is represented on the bar graph by the lengths of the bars, which are proportional to the numbers to be shown.

To introduce comparison of magnitudes, ask pupils to arrange the items in order of decreasing frequencies as shown on either graph.

To give pupils practice in reading scales and comparing magnitudes, have them study such questions as:

- How many more singles did John get than triples?
- The number of home runs exceeded the number of triples by how many?
- How many fewer triples did he make than doubles?

Reading the actual size of each item is not required of the student if he is asked such questions as:

- Would you say John hit twice as many doubles as home runs, or three times as many, or six times as many?
- How many times as many singles as doubles did he hit?
- Would you say John hit one-third as many home runs as doubles, or one-half as many home runs as doubles?
- The number of triples is what fractional part of the number of home runs?
- The number of walks is what fractional part of the number of singles?

These comparisons may be made by estimating the ratios of the bar-lengths.

From reading the title and captions we find that information about total hits is not given on these graphs, but could be found by combining some items. Pupils may be asked such questions as:

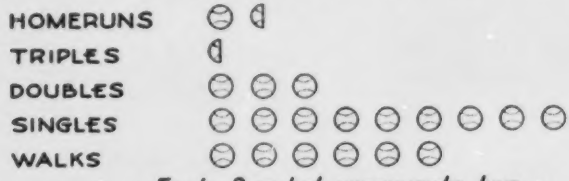
- How many extra-base hits did John have?
- About how many times as many singles as extra-base hits did he have?
- About how many times as many hits as walks?
- What fractional part of John's hits were home runs?

Some people may have enough background to consider related outside factors in answering questions such as: Why were there more home runs than triples?

12. Interpreting Scale Pictographs

In a scale pictograph, each symbol stands for a number other than 1; for example, each might represent 5 automobiles.

1956 Baseball Record of Hits and Walks for John Perkins



Display the scale pictograph shown here and the unit pictograph from activity 11. Ask the pupils to compare titles and captions of the two graphs.

To illustrate the importance of noting the scale used before reading a graph, discuss whether the two graphs show identical information. Have a pupil explain what each symbol on the scale pictograph represents. The class should then count the symbols for each item and tell what magnitude they represent.

To give practice in reading the scale, remove the unit pictograph from view of the class, and ask for comparisons such as:

- How many more symbols are there for singles than for walks?
- If each symbol stands for 2, how many more singles were there than walks?
- How many fewer triples were hit than home runs?

To read and combine items, consider such questions as:

- How many hits did John have?
- How many times did he get on base?
- How many more times did John get on base by hitting than by walking?

It may be pointed out that it is not necessary to use the scale in order to make such comparisons as:

- How many times as many doubles as home runs did John hit?
- The number of doubles is what fractional part of the number of singles?
- What fractional part of the times on base did John get by hits?
- What percent of the times did John get on base by walks?

13. Using Bar Graphs

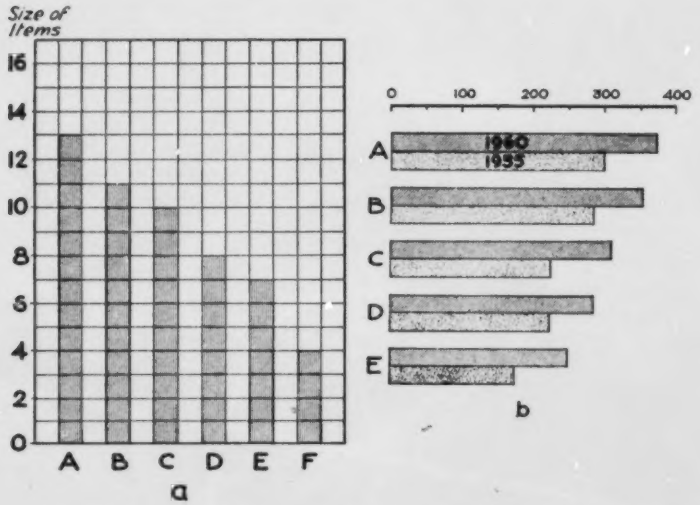
Description: A set of uniformly placed horizontal or vertical bars of equal widths, each starting at the zero point of a scale placed parallel to the bars. Spacing between bars should be uniform, unless the spaces represent intervals of time. In such cases, each space should be pro-

portional to the time elapsed. Zero should be included on the scale which is used. The title and the graph must agree. The identity of each bar should be clear. (See *a.*)

Uses: To permit immediate comparison of quantities or sizes. To emphasize the difference in amounts rather than the fluctuation of something over a period of time. (May be used to show the fluctuation of a sequence consisting of only a few values.)

Interpretation: Conclusions may be drawn from the amounts or percents represented by comparing the lengths of the bars.

Variations: A *Grouped Bar Graph* consists of two or more bar graphs on the same set of axes. Related bars are placed together (grouped) in order to simplify interpretation. (See *b.*)

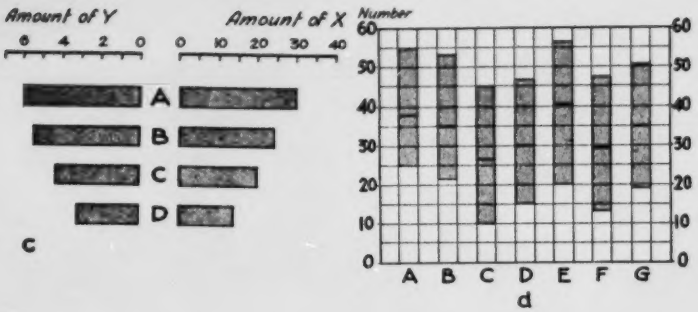


A *Paired Bar Graph* consists of two bar graphs placed so that each bar label identifies two (paired) bars. Separate scales extend on either side of the labels to indicate the lengths of the bars. (See *c.*)

A *Sliding Bar Graph* consists of a set of staggered bars, each of which represents the range[1] for that item. Some value, such as an average, may also be identified along the bar by an appropriate symbol or mark. (See *d.*)

Materials for the Construction of Bar Graphs. Colored plastic tape can be used for making colorful bars in a bar graph, or, if wider bars are to be used, they can be cut from sheets of colored construction paper.

[1] For explanation of "range" see activity 45.



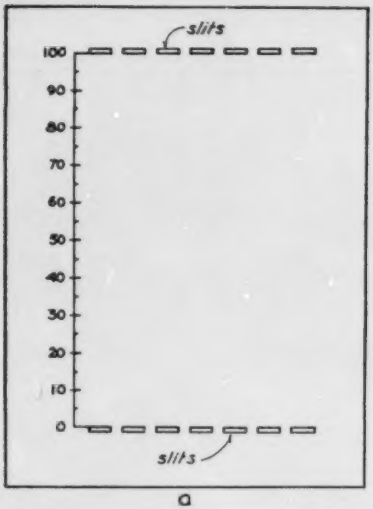
14. Making an Adjustable Demonstration Bar Graph

Make a demonstration bar graph, using a piece of matboard 18″ x 24″, several pieces of brightly colored 1-inch ribbon, and an equal number of pieces of white 1-inch ribbon—each 20 inches long. See page 20.

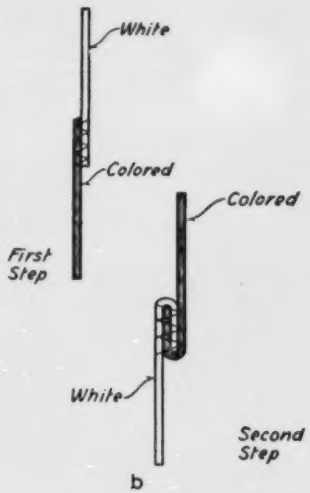
Draw a line parallel to and 3½ inches from the left and long side of the matboard. Starting at a point on this line 4 inches from the bottom edge of the matboard mark ⅞-inch divisions. At the top and bottom of the scale cut a series of seven 1¼-inch slits parallel to the bottom edge of the matboard. These slits should be wide enough to allow the ribbon to pass through easily and be spaced ½ inch apart as shown in *a*.

Glue one end of a colored ribbon strip to an end of a white ribbon strip or sew the ends together to make a French seam as shown in figure *b*. Thread the ribbon through two slits, one directly above the other, pull it taut and fasten the free ends together to form a "belt." If the ribbon cannot freely move up and down, insert a small folded card in each slit for the ribbon to slide over. Do the same with the remaining pieces of ribbon.

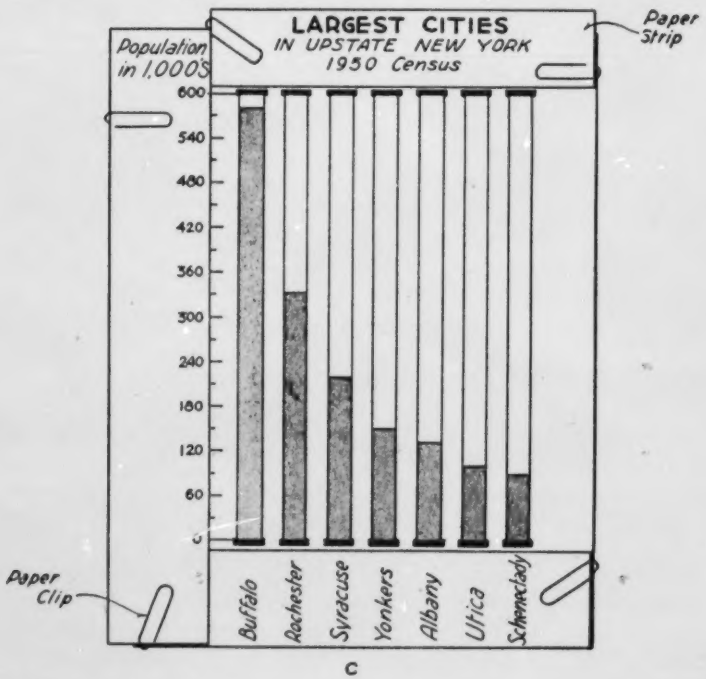
These colored "belts" can now be used as adjustable bars of a bar graph. Strips of paper can be fastened with paper clips at the top, bottom and left side as needed for the title and other information for a specific graph. These strips are easily changed to illustrate a different graph. If the given scale does not apply for a particular problem, cover it with the side-strip of paper and add numerals for another scale. Figure *c* shows a complete graph with label-strips added. This device can stand on a chalk-tray or be fastened to a bulletin board for display. It it is placed on a chalk-tray, added information may be written on the chalkboard.


slits
100
90
80
70
60
50
40
30
20
10
0
slits

a

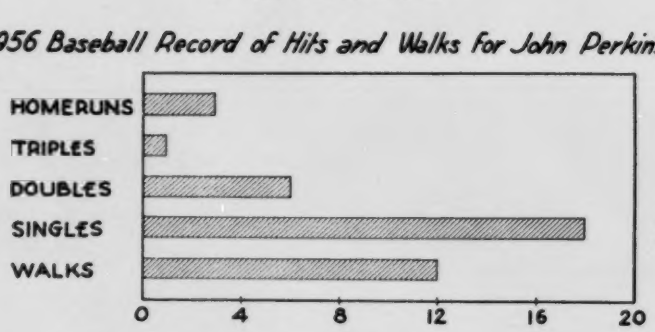

White
Colored
Colored
First Step
White
Second Step

b


LARGEST CITIES
IN UPSTATE NEW YORK
1950 Census
Paper Strip
Population in 1,000's
600
540
480
420
360
300
240
180
120
60
0
Paper Clip
Buffalo
Rochester
Syracuse
Yonkers
Albany
Utica
Schenectady

c

20



15. Demonstrating Interpretation of Scale Bar Graphs

Display in the classroom a scale bar graph, such as that shown in the figure, of sufficient size so that the title, scale and key can easily be read by all pupils.

To identify the subject matter of the graph, have a student read the complete title and captions, giving examples both of what *can* be found on this graph, and what cannot be found on it. It may be discovered that the title of the graph is misleading.

Ask another student to read the scale numbers and to point out any irregularities which appear.

Consider separately each item and the bar representing it. Hold a yardstick, or project an imaginary straight line, at right angles to the bar at its end; read the scale divisions between which it falls. Estimate the fractional part of the distance from the lower scale division to the higher, which is cut off by the projected line. Compute that part of the difference between the two scale values, and add it to the smaller scale value.

For example: in the figure, the length of the bar representing singles is between 16 and 20—about halfway between them. The difference between 16 and 20 is 4; half this is 2. The number of singles is 16 + 2, or 18.

To have pupils practice reading the scale, ask for comparisons telling "how many more" or "how many less"; this will require the reading of two or more items.

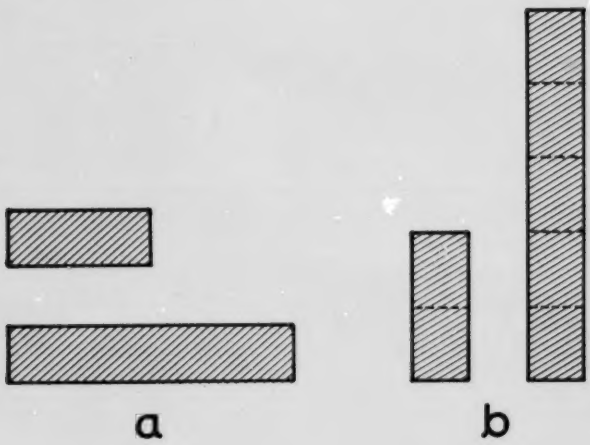
Notice that reading the scale is not necessary for comparisons which involve *order of size, how many times as large* one item is, as compared to another. These comparisons can be approximated without using the scale. To illustrate this, cover the scale by fastening a strip of adding machine tape over the scale temporarily. Ask students if they are still able to compare the magnitudes of the items in any way.

16. Comparing Items in Bar Graphs

To find the ratio of bars in a bar graph, measure the smaller of the two bars to be compared, using a compass, dividers or a tape measure. Then transfer this distance to the larger of the bars, and "step it off" on the larger bar, to see how many times the smaller can be contained in the larger.

In *a*, the ratio of the bars is 2:1 or 1:2. The larger is twice the smaller, or the smaller is one-half the larger, since the smaller can be contained in the larger two times.

In *b*, the ratio of the bars is $2\frac{1}{2}$:1, since the smaller can be contained $2\frac{1}{2}$ times in the larger. By making dashed lines as shown, show that the smaller is $\frac{2}{5}$ of the larger.



Section II

Line Graphs

17. Using Line Graphs

Description: Usually a broken or curved line placed with reference to a set of scaled axes. The vertical scale generally starts at zero and increases from bottom to top. The horizontal scale increases from left to right and is used for the time or some other variable.

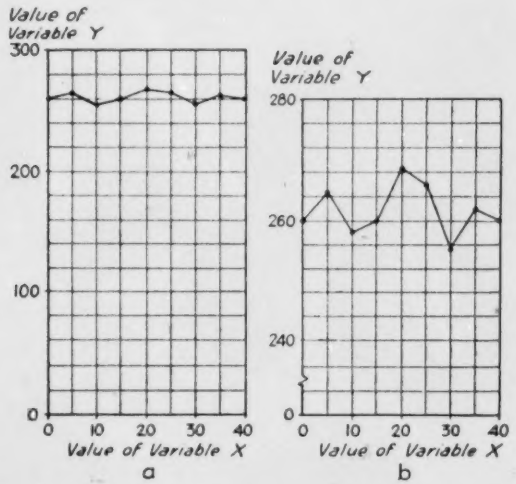
Uses: To show trends or cycles
To plot a series of many successive values
To emphasize movement rather than actual amounts
To identify relationships between variables

Notes: Do not use line graphs to plot a series of only a few values if emphasis is on the difference or change in amounts rather than the movement of the series, or if movement is extremely abrupt or irregular.

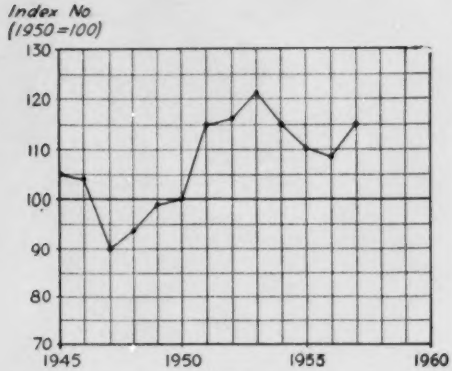
Spacing of the time scale should be regular. Include zero on the vertical scale as in *a*. (Zero can be omitted only when a line other than zero is used for a reference line, as with index number[1] as in *c*.) If too much of the scale is unused, zero should be shown and followed by a break in the scale, as in graph *b*.

Label lines rather than spaces.

Title and graph must agree.



[1] *Index number:* see activity 49.



c

Contracting or expanding the scale can affect greatly the appearance of the graph. Compare *a* and *b*.

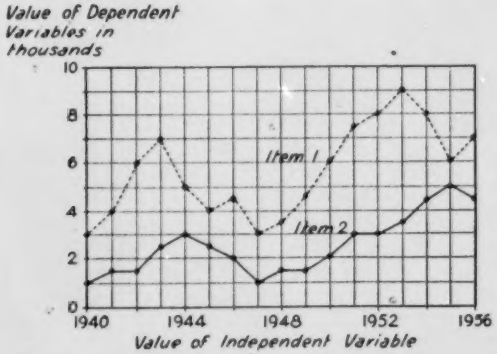
Interpretations:

Comparisons of amounts or percents represented by different points on the line are made by comparing the scale values of these points.

If the vertical scale is unbroken from zero, compare the heights of the points above the horizontal ones.

Variation:

A *multiple line graph* consists of two or more line graphs plotted on the same set of axes for comparison, as in *d*. Each line graph must be differentiated from the others and labeled.



d

18. Making Pegboard Line Graphs

In order to demonstrate line graphs, obtain a piece of pegboard at a hardware or lumber store and paint it with slate-board paint, or any other appropriate paint suggested by the paint store. Different colored pegs, such as are used in a child's peg-bench, can be obtained by the gross from toy supply houses.[1] Have some pieces of different colored strings or elastic thread, which will contrast with the color of the pegboard. Tie metal washers on both ends of each of these strings.

A scale and key can be written directly on the pegboard in white or colored chalk. The lines of holes serve as grid lines. To plot points, put pegs into the appropriate holes, rounding off data to suit the scale. Place a washer on the first peg and stretch the colored string over or under the following pegs, until the last peg is reached. If plain string is used, the remaining string will hang down like a plumb line. If elastic thread of an appropriate length is used, the washer can be placed on the last peg, and the thread will remain taut.

Pegs can be changed or added at will, with very little readjustment of the string.

Multiple line graphs may be represented by using strings of different colors.

Make sure means are available to hang or stand the board up. A large metal stand, an easel or a wide chalk-tray are some possibilities for supporting the board. It may be hung on hooks from the top of the chalkboard. If it is suspended on wires which are not parallel, the board will have less tendency to swing back and forth.

19. Using Line Graphs with a Time Axis: Cycles

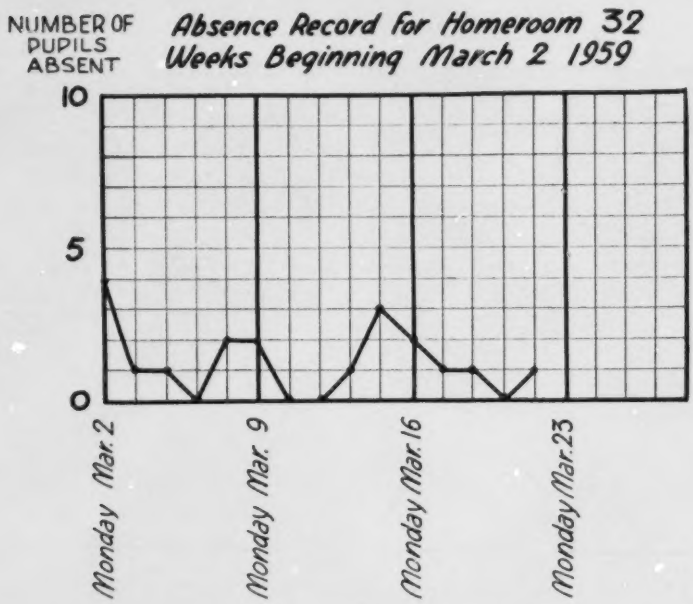
Obtain, or construct, for display a line graph on which one axis is a time axis as in the graph, "Absence Record for Homeroom 32", page 26.

With the class, interpret the title, the time scale, and the number scale with its caption. Read the frequency represented by the point plotted for each day.

Pupils should be cautioned against misinterpretation of line graphs that are obtained from discrete data such as these. The line connecting the points is merely an aid to the eye in following the trend of the data. Numerical readings taken at intermediate points between vertical or horizontal grid lines have no meaning and should not be taken.

A space of time in which a succession of events is completed is called a cycle, if the same succession of events recurs continually.

[1] Source of pegs: Playskool Manufacturing Company, 1750 N. Lawndale Avenue, Chicago, Ill. Item No. 40, Peg Refill.



To illustrate the meaning of "cycle" ask the class to find on the graph the total Monday absences, the total Tuesday absences etc., and to tell on which days of the week the greatest number of absences occurred. Find the average number of absences on each day of the week. Notice that there seems to be a weeklong cycle of absences: on Monday there is usually the greatest number of absences; on Tuesday, Wednesday and Thursday absences are fewer; on Friday, the figure increases again.

Ask the class to predict what the absences might be for the fourth week.

Have students plot similar graphs from data that apply to their own homerooms.

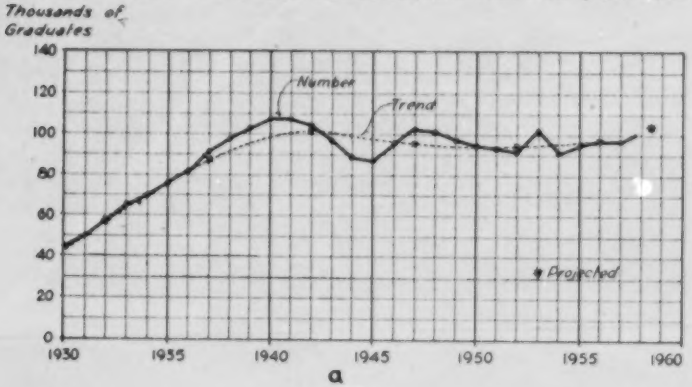
Data which follow a distinct cycle may be best shown by means of a polar graph. (See activity 25.)

20. Using Line Graphs: Trend

Line graphs which have a time axis may show a "trend"; that is, the line shows a prevailing tendency to rise or fall, as times goes on.

Display a graph, such as that shown in *a*, and ask the class to compare the size of each year's graduating class with the size of the previous year's class, telling whether there was an increase or a decrease in size.

NUMBER OF PUBLIC HIGH SCHOOL GRADUATES, NEW YORK STATE, 1930–1958



a

Ask a pupil to estimate from the graph the first five readings, and to give the average for those years. Other class members may give the averages for the other consecutive five-year periods. These averages may be plotted for the third year of each five-year period, and joined to form a trend line, as shown.

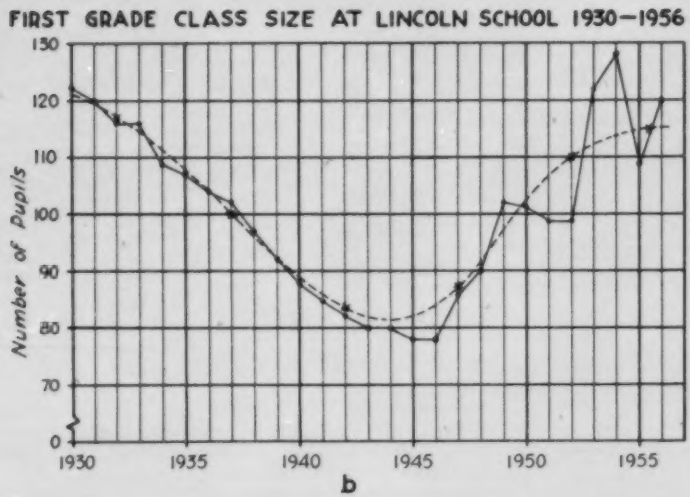
Discuss such questions as:

- Is a general trend noticeable? If so, what is the trend?
- What was the trend between 1930 and 1940? between 1940 and 1950? since 1950?

Show a graph similar to *b* or to any smaller part of *b*. Consider whether there is a single trend noticeable, or whether there is more than one trend. Identify trends and periods for which they were in effect.

Discuss the possible effect of each of the following facts on the size of the classes:

- The depression, from about 1929-39, caused a decline in the birth rate; fewer jobs were available for teenagers and others; unions and others worked to raise to 16 the minimum age for leaving school. See graph *b* on page 28.
- World War II, from 1941-45, brought an increase in marriage and birth rates, more jobs both in service and at home were available for teenagers and others.
- The GI Bill of Rights, which paid the cost of education for veterans, encouraged some to finish high school after returning from service. The effect of this was felt from 1946 to 1955.
- Postwar prosperity from 1948 on has kept the rate of marriages and births high.



21. Interpreting Multiple Line Graphs

Display a mulitple line graph such as that shown. Have a pupil read the title, the captions, the scale and the key.

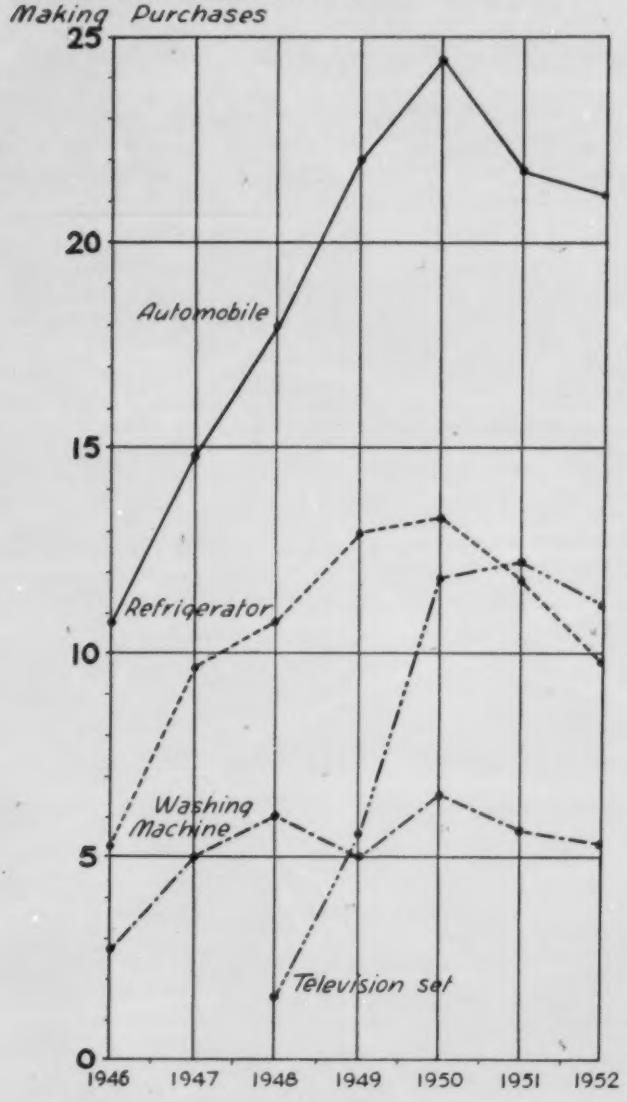
One pupil may choose an appliance and tell what percent of the families in the United States purchased that item in each of the years shown. Is any trend in the purchase of that item evident in the graph? Read the graph for each of the other items, noticing the trends, if any. In three out of four cases, the peak year was 1950, after which there was a leveling off period and sales fell somewhat.

In comparing items on a percent graph, certain cautions are necessary:

- It will not be advisable to compare as to amounts more or less. It would be incorrect to say that approximately 5 percent more families bought automobiles in 1946 than bought refrigerators that year. "Five percent more than" implies that the base of this percentage is the number of families buying refrigerators, which is incorrect.

  It would be correct, but subject to possible misinterpretation to say that the families who bought cars in 1946 exceeded the families who bought refrigerators by 5 percentage points. This implies that the two percentages are of the same base, which actually is the case.

CONSUMER PURCHASES OF DURABLE GOODS

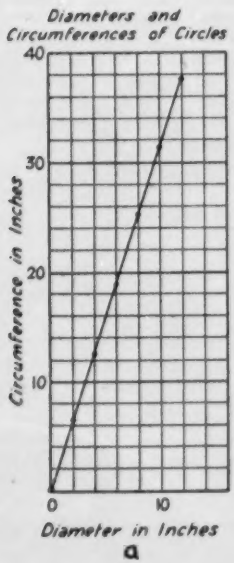


- A better picture can be obtained by comparing these percentages by division: We can say that the number of families who bought automobiles in 1946 was about twice the number who bought refrigerators.
- Neither will it be possible to group items, since the same families may already appear in two or more of the groups pictured. For instance, if this graph included the percentage of families who bought nondurable goods, such as food, the apparent total of all groups would be more than 100 percent of the families in the United States, which is obviously in error.

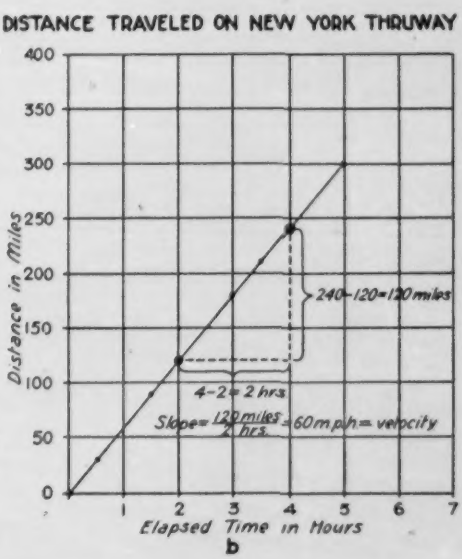
Consider the influence on the graph of such factors as production in World War II and the Korean War; the usual number of years of use of each product before it would need to be replaced; the classification of each item as a necessity or as a luxury.

22. Identifying Relationships between Variables

Display a graph of a functional relationship such as that shown in a. Have a student read the title and captions, and explain what information is contained in the graph.



a



b

Ask a pupil to explain the scales used, and to read the approximate scale values corresponding to the points plotted on the graph. These may be written in tabular form, to help the class see the relationship involved:

Diameter	2"	4"	6"	8"	10"	12"
Approximate Circumference	6.3	12.6	18.8	25.1	31.1	37.7

Compare the numbers in each pair to find the difference between them: point out that the difference is not always the same.

Ask a pupil to tell how many times as big the circumference is, in each case; notice that the circumference is always approximately the same number of times as large as the diameter: a little more than 3 times as large. Thus, the circumference is seen to be directly proportional to the diameter; it varies directly as the diameter.

Ask the pupils to tell other pairs of numbers which would vary directly with each other. Some possibilities are radius and diameter, total cost and number purchased, distance traveled and rate of speed, equivalent measures such as the number of inches in a given number of feet, the number of accidents with the number of automobile miles driven, the incidence of cancer and the number of the population over 40 years of age, or the county automobile insurance rates with the loss due to accidents in the county. Many other illustrations can be pictured in graphic form for the pupils to interpret.

Variation:

Demonstrate that the diameter varies directly with the circumference, being a little less than one-third of the circumference.

Example *b* is a graph of the following data taken by a driver on the New York Thruway.

Distance Traveled on New York Thruway

CLOCK TIME	ELAPSED TIME IN HOURS	DISTANCE IN MILES
8:30	0.0	0
9:00	0.5	30
10:00	1.5	90
11:30	3.0	180
12:00	3.5	210
1:30	5.0	300

Slope:

Data from quantities which vary directly with each other always produce straight line graphs. The slope of such a line is a useful tool in interpreting the graph. The slope of a line is found by dividing the vertical distance between any pair of points on the line by the horizontal distance between the same pair of points.

When graphed the data in *b* show a constant slope and reveal the fact that this driver traveled at a uniform speed of 60 miles per hour.

The use of the slope concept is very helpful in explaining why the arithmetic average is not always a true representation of some other average. For example, if a driver goes from one point to another point at a speed of 20 m.p.h. and then to a third point at the same distance at 30 m.p.h., his average speed is actually 24 m.p.h. rather than the 25 m.p.h. figure that would seem apparent at first glance. Have pupils plot a graph such as *c* to show the truth of this statement.

Data

SPEED DURING INTERVAL IN M.P.H.	LENGTH OF INTERVAL IN HOURS	TOTAL ELAPSED TIME IN HOURS	TOTAL DISTANCE IN MILES
30	2	2	60
20	3	5	120

Slope of $AB = \frac{60 \text{ mi.}}{2 \text{ hr.}} = 30$ m.p.h. = average speed during first two hours

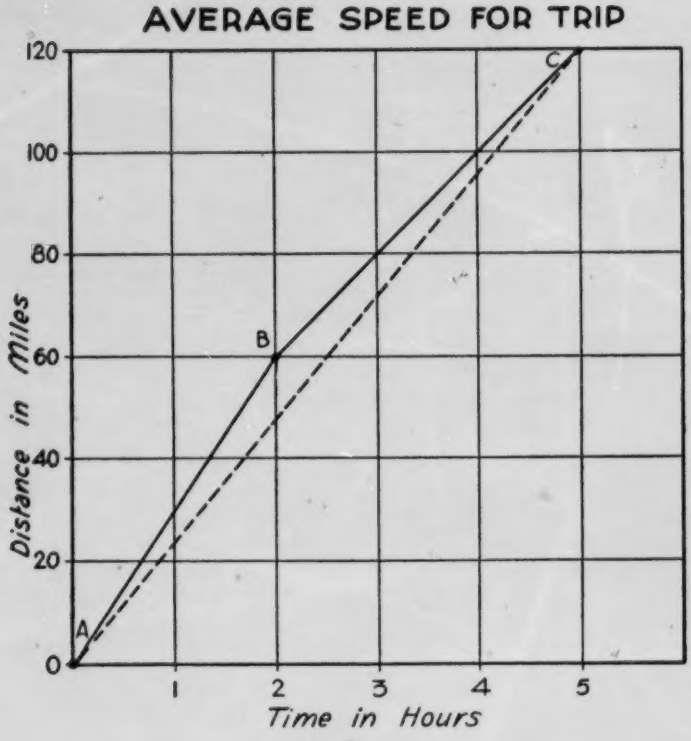
Slope of $BC = \frac{60 \text{ mi.}}{3 \text{ hr.}} = 20$ m.p.h. = average speed during last three hours

Slope of dotted line $AC = \frac{120 \text{ mi.}}{5 \text{ hr.}} = 24$ m.p.h. = average speed for entire trip

Maximums and Minimums:

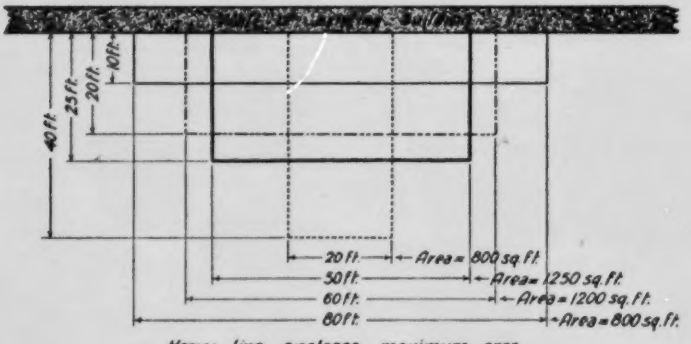
When studying relationships between two variables, it is often found that the value of one passes through either a maximum or a minimum as the other is allowed to increase uniformly. The example cited below shows a maximum.

Have students plot a graph showing the relationship between the area of a rectangular enclosure and its width under the following conditions. Assume a man wishes to enclose a rectangular yard of maximum area along the side of an existing building using 100 feet of fencing. As indicated in *d*, a width of 10 feet produces a long narrow yard of 800 square feet area and extending the width to 20 feet increases the area to 1,200 square feet. However, continued extension of the width does not



c

RECTANGULAR AREAS, THREE SIDES ENCLOSED WITH 100 FEET OF FENCING



Heavy line encloses maximum area

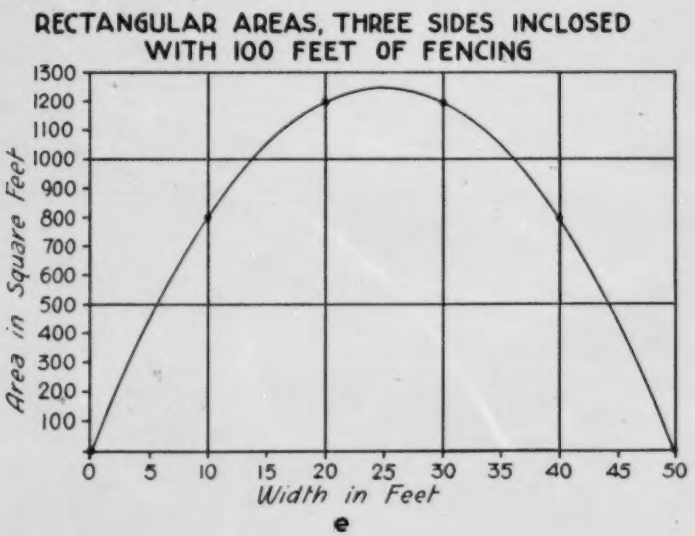
d

necessarily result in additional area. Computed areas for various widths are tabulated below and plotted in *e*.

WIDTH IN FT.	LENGTH IN FT.	AREA IN SQ. FT.
0	100	0
10	80	800
20	60	1,200
30	40	1,200
40	20	800
50	0	0

Since widths of 20 feet and 30 feet both yield areas of 1,200 square feet, it is evident, especially when looking at the graph, that a maximum occurs at the 25-foot width.

Students might also plot graphs of the total profit obtained from the sale of a hypothetical constant-cost item against the selling price under an assumed gradual decline in sales volume as the selling price goes up. This should show a maximum. A minimum can be displayed by plotting a graph of the number of square inches of sheet metal needed to construct a cylindrical can to hold a stated volume. Graph area against diameter and a minimum will be revealed at the optimum diameter. The length of the shadow cast by an object in the sunshine will also reveal a minimum if the data cover time intervals just before and just after noon.



e

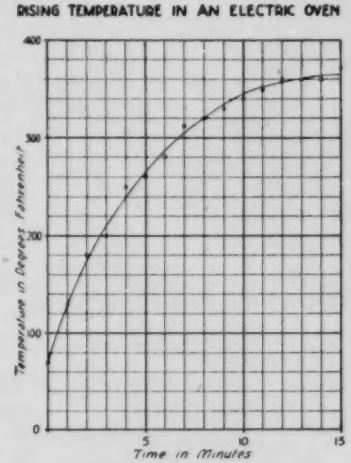
23. Graphing Curves

Expose students to data which form a curve when graphed, lest students assume that all line graphs consist only of points connected by straight lines.

Continuous and varying data should be considered. A number of consecutive temperature readings made at relatively short intervals will provide such information. A local weatherman may furnish the data and perhaps be invited into the classroom to discuss the manner in which he gathers and plots temperature data.

Rising Temperature in an Electric Oven

CLOCK TIME	ELAPSED TIME IN MINUTES	TEMPERATURE IN DEGREES FAHRENHEIT
2:15	0	70
2:16	1	130
2:17	2	180
2:18	3	200
2:19	4	250
2:20	5	260
2:21	6	280
2:22	7	310
2:23	8	320
2:24	9	330
2:25	10	340
2:26	11	350
2:27	12	360
2:28	13	360
2:29	14	360
2:30	15	370



The temperature of an electric oven was recorded every minute for a 15-minute interval after the oven had been turned on with a thermostat setting of 375 degrees. The thermometer used was accurate only to plus or minus 10 degrees. Note in the graph that if each pair of points were connected with a straight line, the resulting graph would be a jagged line. Such a graph would incorrectly imply that the rate of temperature increase changes suddenly at the end of each one-minute interval. It should be pointed out that these apparent fluctuations in rate of heating are due to inaccuracies in the thermometer and that a much more accurate representation of the facts is shown by drawing a smooth curve.

Another source of continuous data is the electric meter found in most homes. Readings can be taken at relatively short time intervals during

the day to show the increase and decrease of electrical power usage. These data can be expected to yield a curve having dips, rises, and flat portions (plateaus).

Cooldown curves of liquids, such as oil when exposed to very cold temperatures, also provide interesting illustrations. If the school has a well-equipped science department, the science instructor may be able to set up a demonstration, using a container of oil immersed in supercooled brine solution. Careful readings from a thermometer will provide data for a curve.

Students should have some practice in graphing data such as are suggested above. They should have the following precautions in mind when drawing such graphs.

- The graph must have a title.
- The graph should show the date on which data were taken unless this information is self-evident from the nature of the plotted data.
- Each axis should be identified with a suitable label.
- Axis labels must answer two questions; i.e., what is plotted, and in what physical units.
- Horizontal and vertical scales must be chosen with these principles in mind:
    1. All points represented in the data should be included.
    2. For the sake of convenience in plotting data, the choice of numerical values for principal grid lines should be simple multiples of the number of subintervals between these grids.
    3. Values for principal grid lines should be chosen so that the graph covers as much as possible of the entire piece of paper on which it is plotted.
- Show plotted points with a circle, large dot or other symbol large enough to be clearly visible after the line is drawn.
- Graphs of continuously varying quantities should be smooth curves. Due to inevitable inaccuracies of measurement, the plotted points usually fail to follow a perfectly smooth curve. As shown in example *b*, the curve should follow the general direction indicated by the points but does not necessarily touch every point.
- Avoid the use of "homemade" graph paper which students are prone to construct from ordinary lined notebook paper.
- Lines on the graph should be drawn with either a straight-edge or a French curve.
- On multiple-line graphs each line should be drawn using distinctive symbols or colors which are identified by means of an appropriate key.

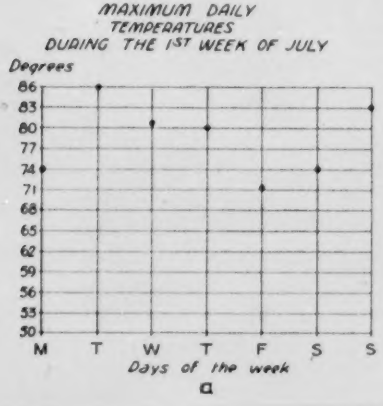
24. Graphing Discontinuous Quantities

Students who become familiar with plotting points on a graph and connecting them to form a line are likely to conclude that all graphs which contain plotted points should also contain connecting lines. A line graph such as one representing a child's growth over a period of years obviously implies that the line connecting any two points is an approximation of the growth over that interval of time. This concept is certainly correct, but the student should not generalize and conclude that all points should be connected with a line approximating the changes which occurred within the interval.

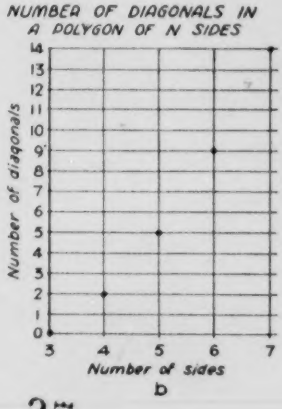
Some sets of data are not continuous; between two consecutive members of a sequence there may be no other members of the same sequence. For example, if the highest daily temperatures were recorded for a period of a week and plotted on a graph (as in *a*), there would be no other temperatures than those indicated. Therefore, a line connecting the plotted points would have no meaning. Data which are discontinuous are said to be "discrete." Point out to the student that discrete data are most readily presented by means of bar graphs or pictograms.

Further examples of data of this type are:

- The number of diagonals in polygons with different numbers of sides. (See *b*.)
- The number of cuts required to obtain logs one foot long from logs three feet long. It would require, for example, one cut to provide one such log, two cuts for two logs, two cuts for three logs etc.
- The number of delegates to a yearly convention over a period of 10 years, or the number of students admitted to a college for each of several consecutive years, or yearly batting averages of a major league player.



a



b

25. Using Line Graphs in Polar Coordinates

Some data are best presented by means of circular graphs. On such a graph data which conform to a rather rigid cycle may be effectively shown. The information is plotted in polar coordinates upon a graph similar to that in the chart. The concentric circles represent uniform distances from the center. If the circles are evenly divided into sectors and each sector is designated as a certain period of time, information may be readily shown.

In the accompanying table and graph "normal" monthly temperatures for both Albany and San Francisco are given. The cycle of monthly temperatures can be neatly presented in such a graph, and conclusions regarding the seasonal differences and general climates can be effectively drawn.

The teacher may wish to use data such as the following in order to plot additional polar graphs: average monthly precipitation, span of daylight for the first day of every month, or cyclic data over a 24-hour or weekly period. Most of these data may be obtained from an almanac.

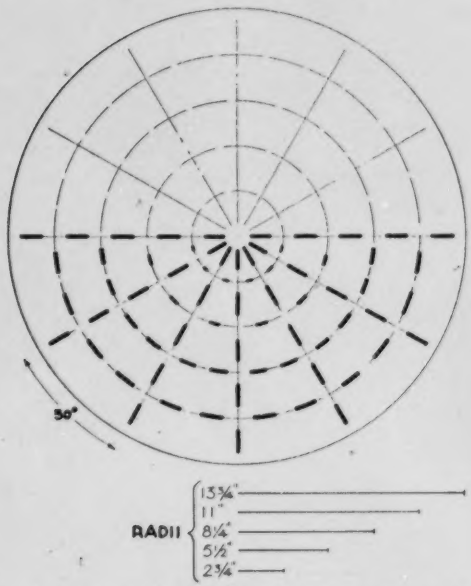
Monthly Normal Temperatures
Albany and San Francisco

MONTH	ALBANY	SAN FRANCISCO
January	26	50
February	27	53
March	36	55
April	47	56
May	58	55
June	69	58
July	72	58
August	70	58
September	63	61
October	52	61
November	42	57
December	31	51

26. Making a Stencil for Polar Coordinate Graphs

In order to simplify the construction of polar coordinate graphs on the chalkboard, a stencil may be cut for use with chalkdust. The stencil should be made from a sheet of firm yet easily worked material. A sheet of six-ply white cardboard, 22″ x 28″, has been found quite satisfactory. From such a sheet a stencil which forms one-half of a polar graph may be cut. A complete graph may be made on the chalkboard by merely turning the stencil over.

GENERAL FORM OF STENCIL FOR POLAR GRAPHS



The polar graph should be accurately drawn on the cardboard with a sharp pencil. Semicircles of radii 13¾", 11", 8¼", 5½" and 2¾" will make a graph which is easily seen by all class members. These semicircles should be made in single pencil lines. The semicircles should be broken into sectors of 30° each. The lines radiating from the center, or spokes, marking these divisions form the basis of the cutouts. Using a line as the long axis, measure slots 3/16" wide and 1½" long, evenly distributed, one between each arc. The arcs should be indicated by curved slots 3/16" wide based on the penciled arcs. One or two slots, depending on the length of the arc between spokes, should be evenly spaced between the spokes. The lengths of the curved slots can be varied between 1" and 1½" in order to permit a proper distribution and leave ample material between the slots to retain rigidity of the material.

The slots may be cut with a razor-sharp knife such as those used in craftwork or with a narrow single-edged razor blade. The curved slots may be cut by fastening the blade to the end of a board compass. *Caution: The cutting should be done only by a responsible adult.*

The completed stencil should be held or taped against the chalkboard and a dusty chalk eraser tapped along the slots to form the outline of a

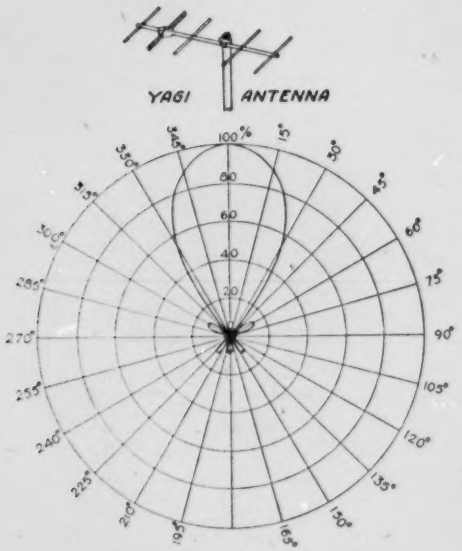
polar graph. The lines may be made to appear sharper if the board is first wiped with a damp cloth.

The figure shows the completed graph as it will appear on the board. As noted previously, however, only one-half of the graph need be cut in the stencil.

27. Indicating Areas on Polar Graphs

Demonstrate that an area which has particular meaning in relation to a fixed point may be effectively shown with a polar graph.

The portion of a road which is effectively lighted by a single automobile headlight may be shown on a polar graph. The limits of the well-lighted area may be indicated by a closed curve. Furthermore, the intensity of light over the surface may be indicated by shading. Sufficient data can be gathered experimentally. One headlight of a parked car can be masked over and a subject asked to approach the single light beam from different angles until he is visible from the driver's seat. In each case the point of first visibility should be recorded in terms of angle and distance.



RECEPTION PATTERN
OF YAGI ANTENNA

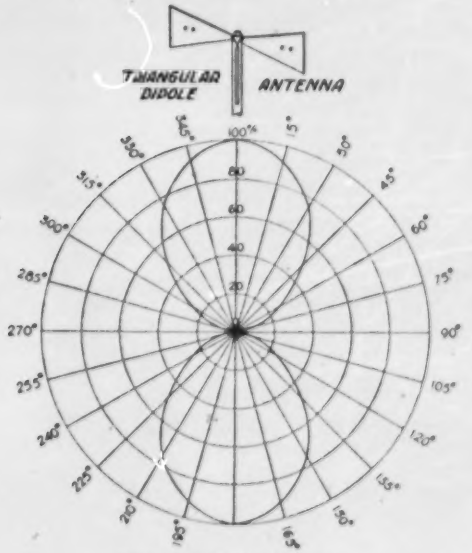
a

Radio and television antennas provide another example of the use of polar graphs. Different types of UHF antennas are used to obtain the best reception in different localities. The reception pattern varies with the type of antenna used. In figures *a* and *b* two familiar antennas are shown together with their reception pattern at a frequency of 500 mc.

The Yagi antenna shown in *a* has a directional pattern which receives best (100 percent of its capability) in only one direction. To the technician this graph shows that it has a rather narrow band of sensitivity which means that it will reject noise and interference.

The triangular dipole pattern has equal sensitivity in two wide directions, which means that it is an effective, yet simple, antenna for areas which have strong UHF signals.

As an exercise to familiarize students with this type of graph, the teacher may have the class consider the problem of picturing the areas which are obscured from a driver's view in daylight because of doorposts and roof supports of his automobile. The center of the polar graph can be considered the driver and his field of vision would be 360°. The students should determine accurately the positions of the obstructions within a specific car and plot them accurately on the graph, then shade in the areas which would not be visible to the driver.



RECEPTION PATTERN
OF TRIANGULAR ANTENNA

b

Section III

Circle Graphs and Composite Graphs

Circle Graph (Pie Chart)

Description: A circle which represents the whole or 100 percent, divided by radii into sectors which represent component parts of the whole.

Use: To show the relationship of relative sizes of component parts to the whole and to each other.

Notes: What the whole circle represents should be clearly indicated.

The sectors are usually arranged clockwise in descending order of size with the largest at the top as shown in illustration. If several small classifications are combined as "miscellaneous" or "other," this sector is placed last even though the total may be greater than other sectors.

Each sector is labeled and usually the percentage or amount represented by it is also indicated.

Percentages and fractional parts of the whole need to be changed to corresponding parts of 360° to determine the central angle for each sector.

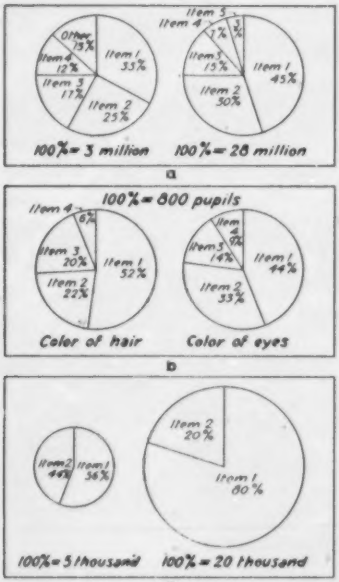
Title and graph must agree.

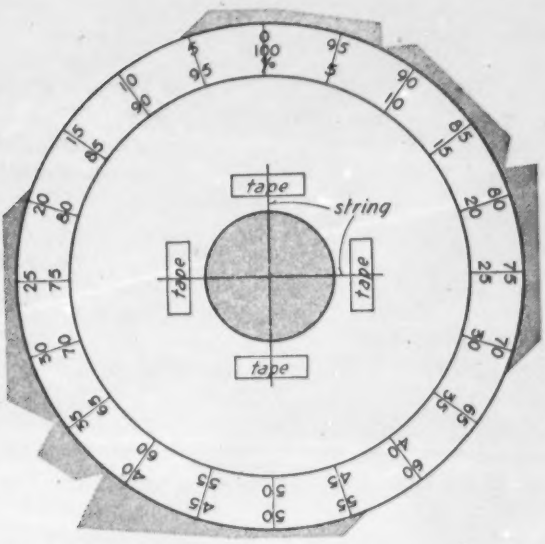
Interpretation: Comparisons of amounts represented by the areas of two sectors, which are not specifically labeled, is made by comparing the lengths of their arcs or the sizes of their central angles.

Variation: A Multiple Circle Graph consists of two or more circle graphs placed near each other for comparative purposes.

Interpretation: The same size circles may represent 100 percent or the whole of two different amounts; equal sections then would represent only equal percents or fractional parts of their respective circle as in *a*. Actual amounts, therefore, cannot be compared until a further computation involving the total or whole amounts is made. Many people misinterpret this graph; they consider that the area of a sector shows an amount rather than only a percent.

The same size circles in another graph may represent equal amounts; equal sectors then do represent equal amounts and can be compared quite easily as in *b*. If each circle represents a different whole amount, the use of proportional circles should show this difference. Their size should be determined by applying the principle that the areas of two circles vary directly as the squares of their radii as in *c*. The area of similar sectors (with equal central angles) would then have the same ratio as the whole circles; thus if two circles represent whole amounts, one of which is four times as large as the other (radius would be twice as large), similar sectors would also represent part amounts, one of which is four times as large as the other. If the sectors are not similar, it is difficult to compare their areas and in turn the amounts represented by their areas.



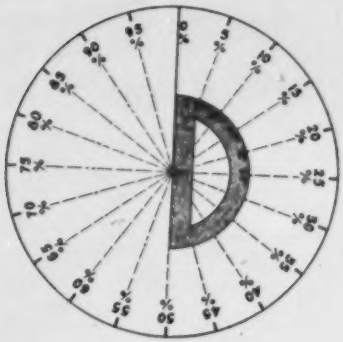


28. Making and Using Percent Protractor

On heavy cardboard or manila folder draw three concentric circles of radii 5", 4¼" and ¾". As in the diagram, draw two diameters perpendicular to each other; using a protractor divide each quadrant into five equal sections (18°); label each of these division lines, which indicate 5 percent. Cut out the 5" and the ¾" circles; tape two strings in line with the perpendicular diameters across the small opening so that they intersect at the center.

This percent protractor can be used at the board to make sectors of a circle graph quickly. Knowing the percentage to be represented and having the circle with its center drawn on the board, draw any radius of the circle and place the cross-hairs of the protractor over the circle's center and the zero percent mark in line with the radius; then going in either a clockwise or counterclockwise direction and using the correct scale on the protractor mark the percent desired, estimating when necessary. Remove the protractor and draw the radius of the circle through this marked point to obtain the desired sector. Replace the protractor with cross-hairs at the center and the zero percent mark in line with the radius just drawn, and continue to mark the next sector.

Variation: Pupils may make a 3" protractor to be used at their desks. This may also be used to read the percentage represented by a sector if percentage is not shown.



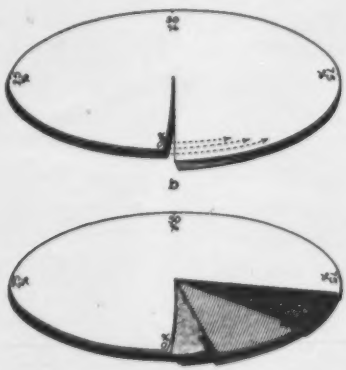
a

29. Making an Adjustable Circle Graph

Cut five different colored circles 12″ in diameter from lightweight poster board. The number of circles used determines the maximum number of sectors that can be represented in the circle graph.

As shown in *a*, use a protractor to mark each circle every 5 percent (18°) from zero percent to 95 percent, in a clockwise direction around its circumference. Draw a single radius to the zero percent mark. Cut each circle out and make a cut to its center along the marked radius.

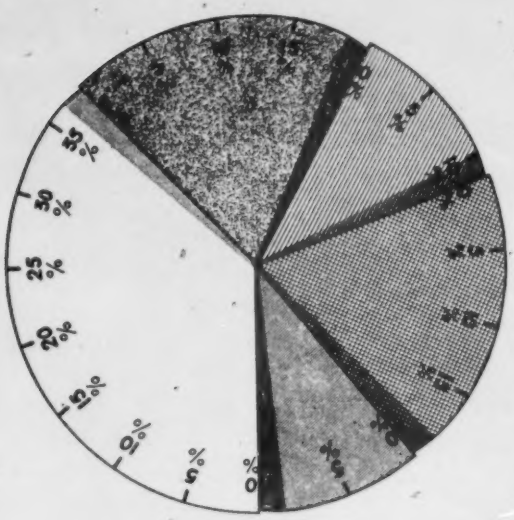
As shown in *b*, stack the circles, matching the cuts. Lift the zero to 25 percent quadrant of the stack of circles, and rotate the bottom circle so that most of its zero to 25 percent quadrant overlaps the top circle, as indicted by the arrows in the illustration. Rotate each of the circles, from the bottom circle up, so that a fanlike arrangement is obtained, as shown in *c*.



b

c

15



d

The final graph is shown in *d* with the percentages of overlap adjusted in accordance with the sizes of the sectors to be represented in the circle graph.

If two or more models are made, they can be used simultaneously to present double or multiple graphs.

If materials are available, have pupils make a similar 4" model. The circles, with percents and radii marked, ready to be cut, can be duplicated on stiff paper. Responsible students may be asked to cut these out, and distribute them so that each pupil has four or five different colors from which to make his adjustable circle graph.

On the reverse of each circle, make a 360° protractor, marking every 5° from 0 to 360° around the circumference of the circle in a clockwise direction. This scale should be marked after the circle is cut out, so that 0° mark may be exactly on the cut.

30. Using the Adjustable Circle Graph Model

Select data, such as the population figures shown below, that lend themselves to a part-whole relationship, from *The World Almanac* or any other source. Round off the numbers and compute the percentage of the total population represented by each continent. Decide on a color key, using available colors on the graph model and adjust colored sectors of the model to show the computed percentages. Paper clips can keep

the size of each sector from changing until the graph is completed. The following table is a convenient way to organize the data to be used.

CONTINENT	ROUNDED POPULATION	PERCENT	COLOR KEY
Africa	208,000,000	9%	Dark Blue
North America	230,000,000	10%	Red
South America	118,000,000	5%	Light Blue
Asia (except USSR)	1,307,000,000	58%	Yellow
Europe (except USSR)	403,000,000	18%	Flesh
	2,266,000,000	100%	

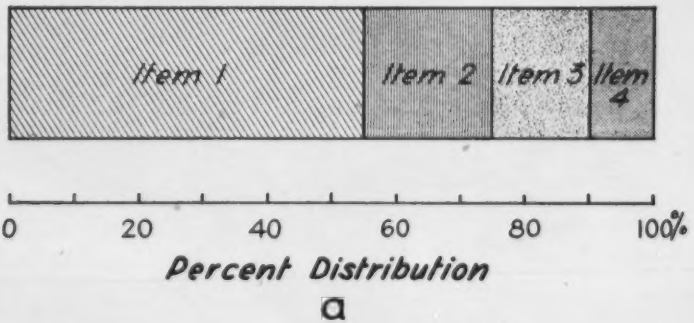
Variation: Data of these types can also be pictured on the demonstration model of an adjustable composite bar graph described in activity 32. Starting with the inner sleeve, each should be extended to the proper scale mark in order to indicate a percent of the whole column; the combined sleeves must represent 100 percent. If fewer than five colors are needed, the colored sleeves not used can be telescoped into one of the longer sleeves.

31. Making a Composite Bar Graph

Description: A single horizontal or vertical bar which represents the whole or 100 percent, which is divided into segments of the same width. Each segment represents a component part of the whole. A scale may be placed parallel to the length of the bar as shown in *a*.

Use: To show the relationship of component parts to the whole and to each other.

Note: What the whole bar represents should be clearly indicated. The segments are usually arranged according to size with the largest at the bottom or to the left. If several classifications are combined as "other"



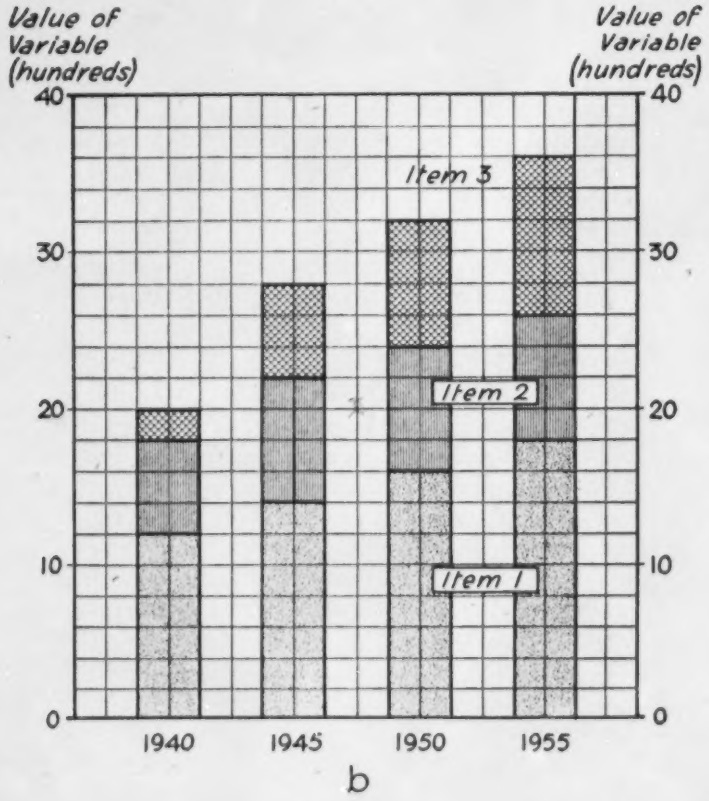
a

or "miscellaneous," this segment is placed last even though its total may be greater than other sectors.

Each segment is labeled and usually the percentage or amount that it represents is also indicated.

Title and graph must agree.

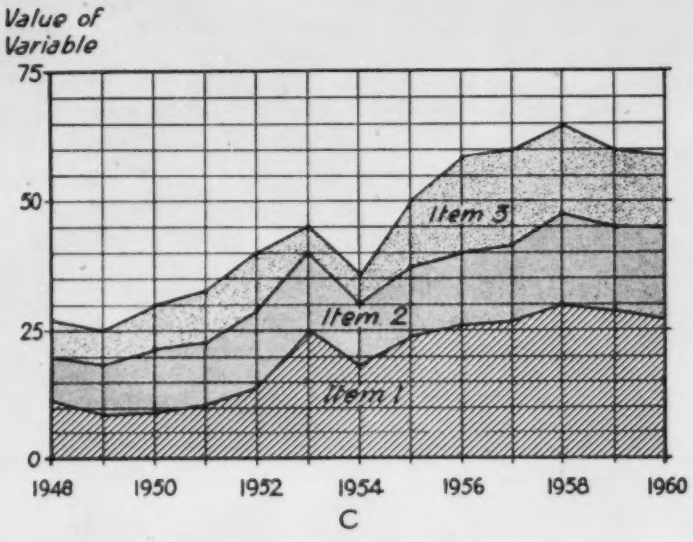
Interpretation: Comparisons of amounts represented by segments which are not specifically labeled are made by comparing the lengths of the segments since their widths should be the same.



b

Variation: A Multiple Composite Bar Graph consists of two or more composite bar graphs which use the same scale and are placed parallel to each other for comparison as shown in *b*.

A Band Graph is a multiple composite bar graph placed without spacing between bars. Corresponding segment division marks are connected with a broken or curved line so that it appears to be a multiple-line graph as shown in *c*. It must be read as a composite bar graph.



C

32. Making a Demonstration Model of an Adjustable Composite Bar Graph

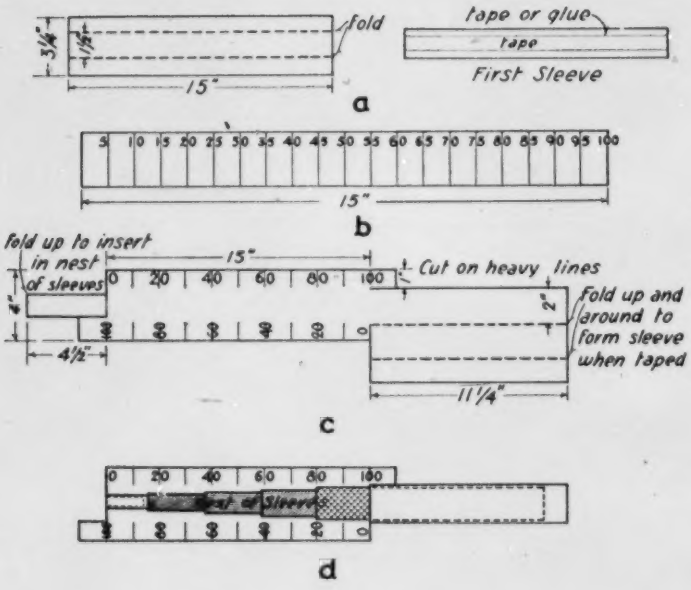
An adjustable composite bar graph can be made, using a set of colored paper sleeves that fit inside each other in telescope fashion.

Using paper of manila-folder weight, make a flat sleeve 1½″ wide and 15″ long. It will require a piece 3¼″ wide to allow for lapping and taping as in *a*. Using this inner sleeve as a guide and a different colored paper cut ¼″ wider than the first piece, make a second sleeve by folding it around the first and taping it. This sleeve should be able to slide over the first easily but still fit snugly. Continue to make sleeves, using the preceding one as a guide, until one has been made for each different color to be used. (Five colors should be enough for most graphs.)

Mark and label each sleeve as in *b*, using ¾″ spaces between division lines. Leave the inner sleeve as it is; cut the second sleeve along the 95-line; the third along the 90-line; the fourth along the 85-line and the fifth along the 80-line so that when nested together each sleeve will ex-

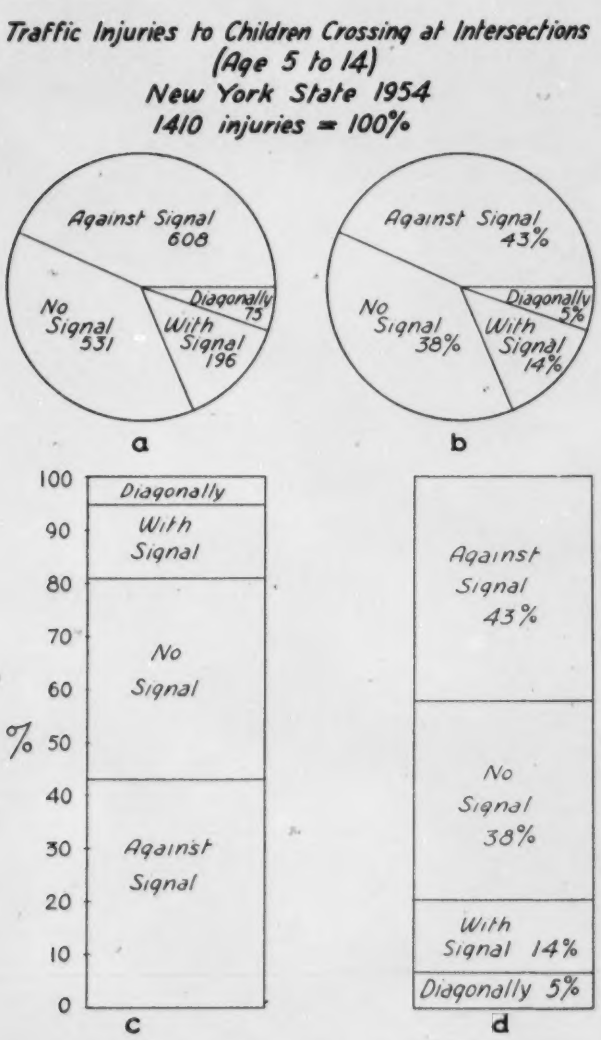
tend ¾" beyond the next outer one. This enables them to be separated easily and avoids one getting lost within another.

Next mark a pattern as in *c* on white paper of the same weight as the sleeves. This pattern when cut out on the heavy lines, folded on the dotted lines and taped will form a holder for the sleeves as in *d*. The numbered scales on the holder and sleeves indicate percent. When inserted in the holder and extended, the visible portion of the nest of sleeves includes 100 percent and can be used as a composite bar graph. The two scales of the holder permit it to be used as a horizontal or vertical bar; its sleeve portion is convenient to hold the graph while demonstrating.



33. Interpreting Circle Graphs and Composite Bar Graphs

Before class begins place on chalkboard circle graph as in *b* or any other that will be of interest to the class. If board space is not available, a poster large enough to be read by the class may be used and then stored for future use.



Discuss with the class questions based on the interpretation of this graph such as the following:

- From the title, what would you expect the graph to show?
- How many injuries are included in the total shown?
- How many injuries resulted from crossing diagonally?
- What method of crossing resulted in the most injuries? the least?
- The number of injuries crossing against the signal is approximately how many times as great as the number of injured crossing with the signal?
- The number of injuries crossing with signal is approximately what fractional part of the injuries crossing against signal?
- Approximately what fractional part of the injuries were caused by crossing against the signal?
- Approximately what fractional part of the injuries were incurred by crossing illegally?
- What generalizations (general conclusions) could you make from the graph?
- Explain why the conclusion, "It is safest to cross diagonally at intersections," is not valid.

Show each of the poster-size type-graphs (*a*, *c* and *d* previously prepared) one at a time, and reconsider some of the previous questions, noting any similarities in securing answers.

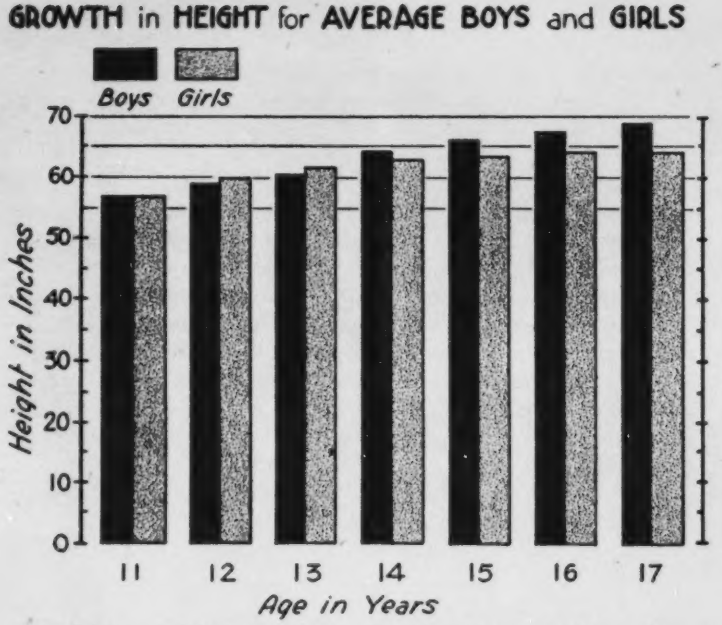
34. Interpreting a Grouped-Bar Graph

A grouped-bar graph shows two sets of data on a single pair of axes. There should be a key to tell what each side of the double bar represents, as in the accompanying example.

Comparisons involving order of size can be made without reference to scale. Ask students to identify:

- The age at which boys' and girls' average heights are equal.
- The age at which the girls are taller.
- The ages at which the boys are taller.
- The age at which there is the greatest difference in heights between boys and girls.
- The age at which there is the greatest increase in the girls' height over the previous year.

To give opportunities to read the scale, ask students to identify the size represented by each bar; to find increases in height over a given period of years; to find how much more the girls' height is than the boys' height, or vice versa, at any given age.



A different technique in reading scales can be practiced by asking a pupil to tell the age at which the average girl is 57" tall, or the age at which the average boy's height reaches 65".

Outside factors influence individual items, and should be taken into account in interpreting, so far as they are known. Discuss with the class such questions as:

- In what grades would the average girl be taller than the average boy?
- If a boy is of average height at age 11, but at 13 and 14 is below average, will he always be below average?
- If a girl grows three inches in height between 12 and 13, will she always grow faster than average?

35. Reading Composite Graphs (Band Graphs)

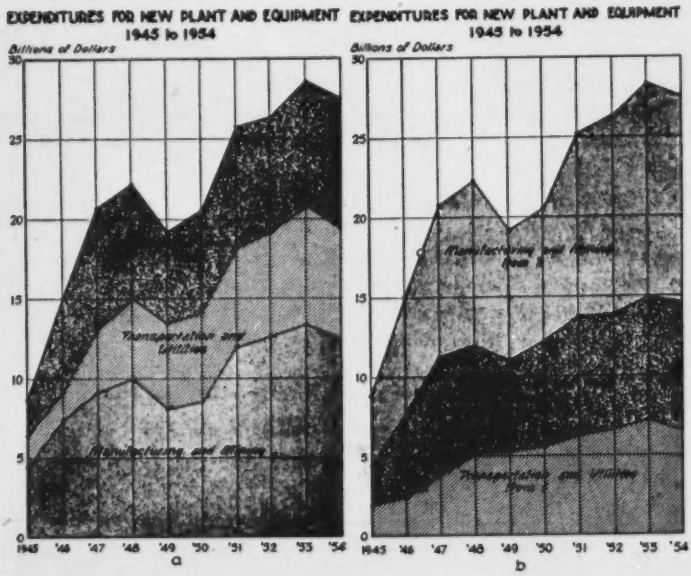
Since band graphs appear in print, and are peculiarly liable to misinterpretation, pupils should be taught how to read them correctly. A band graph appears to be a multiple line graph with shading between the lines. Actually a band graph is a composite graph; the scale values represented by the lines are cumulative values, and the top line represents a total value.

The magnitude of an item on a band graph is represented not by the height of one line, but by the vertical width of the shaded band which represents the item. To find its actual size, read scale values of the intersections of a vertical line with the top and bottom boundaries of a single band, and subtract.

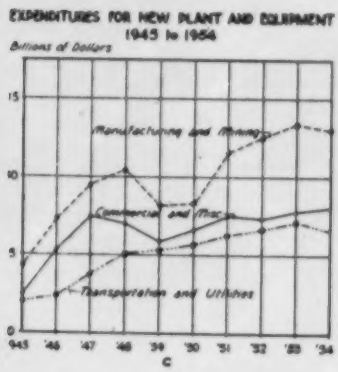
To note the danger of misinterpreting band graphs, study graphs *a* and *b*. The apparent generalization to be made from graph *a* is that there was a general retrenchment in 1949. Graph *b*, which merely changes the order of the bands, gives a somewhat truer picture. Misinterpretations are most likely to center on periods in which there is strong fluctuation in one item, with an apparent correlative fluctuation in other items.

Display the three graphs *a*, *b* and *c*. Have the title, captions, and scale of the graphs explained. Students should be able to read the data for the cumulative totals represented by the lines. These data may be recorded on a duplicated form similar to that shown at *d*. It should be emphasized that it is the *difference* in values which must be determined each time.

Data in form *c* was obtained by reading values for lines *A*, *B*, and *C* on graph *b*, and subtracting adjacent values.



54



Notice that the difficulty in interpolating readings on the scale contributes to the inaccuracy of the resulting data. A better comparison can be made by plotting data for individual items as separate line graphs, as shown in *c*. Readings and comparisons can both be made much more easily.

FORM FOR READING BAND GRAPHS

Year	Line A Item 1 (billions)	Line B Items 1 & 2 combined	(B - A) Item 2	Line C Items 1, 2 & 3 combined	(C - B) Item 3
1945	2	4.5	2.3	8.8	4.3
'46	2.4	7.7	5.3	16.0	7.3
'47	3.8	11.2	7.4	20.8	9.6
'48	5.1	12.0	6.9	22.3	10.3
'49	5.2	11.1	5.9	19.2	8.1
'50	5.6	12.2	6.6	20.5	8.3
'51	6.3	13.8	7.5	25.2	11.4
'52	6.6	14.0	7.4	26.4	12.4
'53	7.2	15.0	7.8	28.3	13.3
'54	6.5	14.6	8.1	27.6	12.9

d

Section IV

Scatter Diagrams

Description: A rectangular or circular graph upon which data are plotted in order to show trends, relationship or the pattern of distribution.

Use: Scatter diagrams are used primarily to show the relationship between two sets of data. On a rectangular graph one set is indicated along the horizontal axis and the other along the vertical axis. Scatter diagrams are also used to give an indication of the pattern formed by a set of objects over a given surface.

Interpretation: Once the data have been plotted on the graph, conclusions may be drawn regarding trends, relationships or limits of the distribution. A major concentration of points in the form of a belt will indicate the presence of a relationship (correlation) between two sets of data. Concentrations, of any shape, are significant in determining the general pattern of a distribution.

36. Showing Presence of Correlation

Use scatter diagrams to show the relationships (or lack of it) between sets of data. If two collections of facts are related, the plotted points will form a distinct concentration. A belt of clustered points, which slopes upward, indicates a significant positive relationship or "correlation"; the narrower the belt the greater the correlation. A perfect correlation would produce a straight line at a 45-degree angle.

In order to develop the concept of correlation and to familiarize students with scatter diagrams, data which are significant and interesting should be used. Statistics such as those gathered in sports, class tests and school surveys are within this category.

The table on page 57 contains information on leading batsmen in the American League for the years 1926-55. Several scatter diagrams can be constructed from these data to detect relationships between the sets of variates.

From graph *a* it is apparent that there is little, if any, correlation between leading batters' home run production and their hitting average. However, graph *b* shows a definite correlation between the number of times at bat and the number of hits obtained for these same men; that is, as the number of times at bat is increased there is a distinct tendency for the number of hits to increase.

It should be noted that the graphs need not start at zero, but they generally provide a range within which all significant points may be

AMERICAN LEAGUE LEADING BATSMEN

1926-55, Inclusive

YEAR	PLAYER	CLUB	G	AB	H	HR	RBI	AVG
1926	Harry Manush	Detroit	136	498	188	14	86	.378
1927	Harry Heilmann	Detroit	141	505	201	14	120	.398
1928	Leon Goslin	Washington	135	456	173	17	102	.379
1929	Lewis Fonseca	Cleveland	148	566	209	6	103	.369
1930	Al Simmons	Philadelphia	138	554	211	36	165	.381
1931	Al Simmons	Philadelphia	128	513	200	22	128	.390
1932	Dale Alexander	Det.-Bost.	124	392	144	8	60	.367
1933	James Foxx	Philadelphia	149	573	204	48	163	.356
1934	Louis Gehrig	New York	154	579	210	49	165	.363
1935	Charles Myer	Washington	151	616	215	5	100	.349
1936	Lucius Appling	Chicago	138	526	204	6	128	.388
1937	Charles Gehringer	Detroit	144	564	209	14	96	.371
1938	James Foxx	Boston	149	565	197	50	175	.349
1939	Joseph Dimaggio	New York	120	462	176	30	126	.381
1940	Joseph Dimaggio	New York	132	408	179	31	133	.352
1941	Theodore Williams	Boston	143	456	185	37	120	.406
1942	Theodore Williams	Boston	150	522	186	36	137	.356
1943	Lucius Appling	Chicago	155	585	192	3	80	.328
1944	Louis Boudreau	Cleveland	150	584	191	3	67	.327
1945	George Sturnweiss	New York	152	632	195	10	64	.309
1946	Mickey Vernon	Washington	148	587	207	8	85	.353
1947	Theodore Williams	Boston	156	528	181	32	114	.343
1948	Theodore Williams	Boston	137	509	188	25	127	.369
1949	George Kell	Detroit	134	522	179	3	59	.343
1950	William Goodman	Boston	110	424	150	4	68	.354
1951	Ferris Fain	Philadelphia	117	425	146	6	57	.344
1952	Ferris Fain	Philadelphia	145	538	176	2	59	.327
1953	Mickey Vernon	Washington	152	608	205	15	115	.337
1954	Roberto Avila	Cleveland	143	555	189	15	67	.341
1955	Albert Kaline	Detroit	152	588	200	27	102	.340

G—Games Played
AB—At Bats
H—Hits
HR—Home runs
RBI—Runs Batted In
AVG—Hitting Average

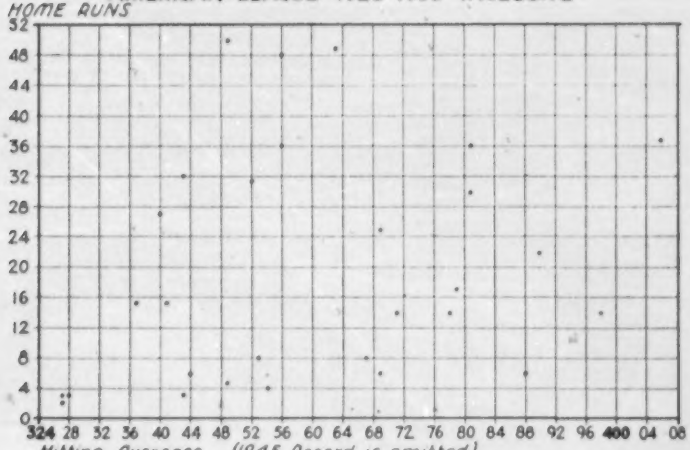
plotted. Moreover, the units on each axis should be generally comparable.

In general, the validity of an observation increases with the number of points which are plotted. Provided they are truly representative, 20 points constituted an approximate minimum. Erroneous conclusions may be drawn from a small number of scattered points simply because the points selected for plotting are less likely to be representative of the whole distribution.

Additional scatter diagrams can be compiled from the table. The comparison of times at bat and number of games played yields a distinct belt, as does the plot of home runs versus runs batted in. When home runs are plotted against number of games played, there is very little correlation.

The teacher may develop other examples of correlation from data gathered within the class. Pupil height plotted against the scores on the last examination should provide an example of no correlation. Com-

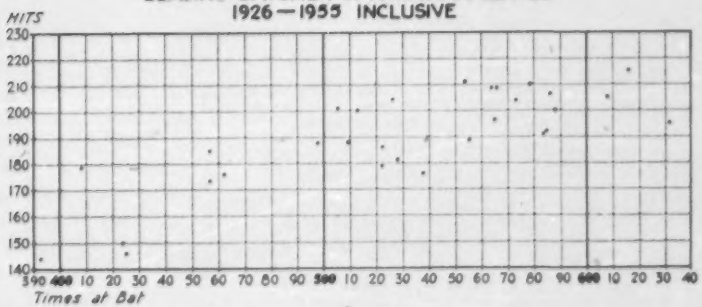
COMPARISON OF NUMBER OF HOME RUNS
WITH HITTING AVERAGES FOR LEADING BATSMEN IN
AMERICAN LEAGUE 1926–1955 INCLUSIVE



a

parison of mathematics and English examination scores for class members will probably indicate some correlation. High correlation will be noted between the heights of students and the heights of their waists (measured at the belt). Very high correlation will be found between the heights of students and their reach; reach being the distance between fingertips when the arms are extended horizontally at the side.

COMPARISON OF NUMBER OF HITS
WITH NUMBER OF TIMES AT BAT FOR
LEADING BATSMEN IN AMERICAN LEAGUE
1926—1955 INCLUSIVE



b

37. Using a Scatter Diagram To Show Pattern: Polar Graph

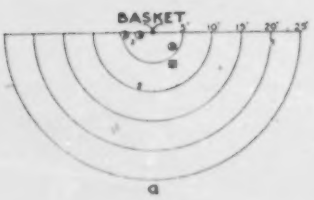
There are some distributions which are best shown by means of polar graphs. Whenever objects are considered with respect to their distance and direction from a fixed point a scatter diagram upon a polar graph is most effective. Polar graphs are described in activity 25.

This type of scatter diagram is commonly used by basketball coaches to gather important information about their own team or about some opponent. During a game the shots taken at the basket by the players on each team are recorded on polar graphs so that the position on the floor and the distance from the basket are indicated. Analysis of the completed graphs enables a coach to determine the most effective and most dangerous scorers of the opposition and their favorite shooting areas, as well as weaknesses in his own team's defense.

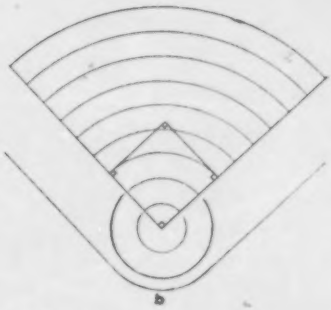
Some students with athletic experience may be familiar with these game charts and will be eager to demonstrate their knowledge in class. The teacher may wish to capitalize on this interest by utilizing these students to explain the use of charts in planning game strategy. Perhaps a word on their use by the coach would add to the interest. Such an approach is fascinating to most students because they feel they are being given real coaching "secrets."

In diagram *a*, a typical game chart is shown. If player number 5 shot at the basket from a point 20 feet away and to the right of the basket a small number 5 was entered at that point. If the basket was made, the number was circled. Note that only half of a polar graph is necessary to record these types of data.

Students can be requested to attend a school game and compile their own game charts. A similar chart showing basketball rebounds may also be kept. Still another use of this type of scatter diagram occurs in base-



a


b

ball, where a record of the hits made by the members of a team is kept on a polar graph. Such a record has led to strategic defensive moves like the so-called "Ted Williams shift." A record of flyballs could be kept on a graph such as that shown in figure *b*.

Polar scatter diagrams are also used to aid in the analysis of other types of distribution. For example, if the holes on a target made by shots from a rifle tend to be clustered in one area away from the center, the shooter knows that his aim or the sights of the gun must be corrected in a certain manner. Likewise a scatter diagram is formed by iron filings sprinkled on a sheet of paper when a magnet is held just beneath. The filings move so as to indicate the lines of force between the two poles. Although this example is somewhat different from the others which have been mentioned, it illustrates that scatter diagrams may vary in type and usage. The teacher may wish to introduce this latter illustration to coincide with its presentation in a general science course.

38. Using a Scatter Diagram To Show Pattern: Rectangular Graph

Recall to students the familiar use of a scatter diagram to show the distribution of a population within a given area. The boundaries of the area are indicated on a graph and the grid lines are identified so that location within the area may be made by means of grid coordinates. The graph is sometimes shaded to indicate the density of population within each section. An example of such graphs may usually be found in an atlas.

MUSEUMS
INCORPORATED BY THE STATE OF NEW YORK

Other illustrations of this type of graph are numerous. The teacher may wish to reproduce an outline graph of New York State and distribute copies to students for use in the development of scatter diagrams. On this outline the location of an object may be indicated by a dot or symbol. Variations, when desirable, may be shown by the use of color or by variations in shape or size of the dot or symbol.

In the accompanying figure, a distribution of museums incorporated within New York State is indicated.

Variation: A roadmap may be attached to the classroom bulletin board and pins with large colored heads placed at points on the map to indicate the locations of a set of objects.

Section V

Statistical Tables

Statistical Tables: Common Form and Variations

The primary object of a table is to present large amounts of data in a precise and orderly manner. An effective table therefore is one which contains a systematic listing of relevant information in a minimum amount of space.

39. Becoming Acquainted with Statistical Tables

In order to assure that students are well aware of the characteristics of good tables and are able to read tables correctly, the teacher should plan to devote a small amount of time to a discussion of tables.

Although tables are individually designed to fit specific material, some general characteristics of effective tables may be noted.

- Information is well-ordered and is systematically arranged (as in columns and rows) so as to provide compact reference data.
- Short descriptive headings and subheadings are used to label each column. Abbreviations are used only where absolutely necessary and only if the meaning is clear.
- Footnotes are used to clarify particular items in the table.
- In the event special directions are required, these appear at the bottom of the table.
- A descriptive title is placed at the top of the table.
- Totals or other subdivisions of the data are set off by means of a double-ruled line or bold horizontal lines.

By far the most common type of table is that which is arranged in columns and rows. A calendar on the classroom wall provides a simple example of this type of table. The students should be familiar with some mathematical tables such as a table of squares and square roots, or a table of measures. To further acquaint pupils with collections of tables, a handbook of mathematical tables may be passed around the class for examination. Students can provide additional examples of columnar tables.

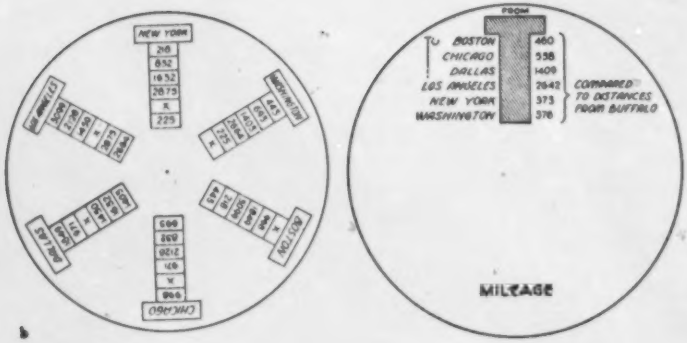
Tables also appear in different forms and it is important that students be aware of these forms. Tables sometimes consist of a single list of information, such as a list of formulas. A table of this type may be found in the review section or appendix of many mathematics texts.

Sometimes a relationship between members of the same group is shown in a table. When this occurs an odd-shaped table results. Consider the mileage table found on most roadmaps, as in *a*. Since the distance between two cities is the same in each direction, one half of the table need not be included because it would only repeat information. This illustrates the use of tables to provide a compact set of ordered data.

	Chicago	Dallas	Los Angeles	New York	Washington
Boston	998	1849	3099	218	443
Chicago		971	2128	832	693
Dallas			1450	1632	1403
Los Angeles				2875	2664
New York					225

MILEAGE TABLE FOR SIX LARGE CITIES

a



b

62

Another type of table is one which is made in the form of a revolving wheel or dial with an aperture that allows different sets of data to appear as the wheel is turned. An example of a circular table is provided in *b*. The disc with the T-shaped aperture is to be centered over the other disc. This table contains the same information as that in *a*, but it is so comprised that the distances from any one city are isolated from the others, and comparison with another set of data may be readily made.

Suggest that class members compete in an effort to discover other unusual types of graphs. Examples which are brought in can be kept in a folder for use in future years.

Sources of Tables

Copies of *The World Almanac* provide much data and many tables.

Governmental agencies (National, State and local) compile and make available numerous tables. The Bureau of Labor Statistics is one such agency. Copies of tables may be obtained through the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C.

The school administrative office is an immediate source of data, as is the school library.

See also suggestions contained in the introduction to section VIII.

40. Reading Statistical Tables

Display a large table such as the one illustrated on World War II Casualties. Pupils may be given smaller duplicated copies of the table.

Have the title read and explained by a pupil. Read the "stub" (the first column at the left) and notice that it further explains the title. Read the column headings and subheadings; discuss which columns were obtained by grouping other columns.

World War II Casualties of U. S. Armed Forces
December 7, 1941—August 31, 1945

BRANCH OF SERVICE	NUMBERS ENGAGED	DEAD AND WOUNDED				
		TOTAL	DEATHS			WOUNDS NOT MORTAL
			TOTAL DEATHS	BATTLE	OTHER	
Army	10,420,000	884,135	318,274	234,874	83,400	565,861
Navy	3,883,520	102,821	65,043	39,379	25,664	37,778
Marines	599,693	91,718	24,511	19,733	4,778	67,207
Total	14,903,213	1,078,674	407,828	293,986	113,842	670,846

Ask the class to identify several numerical items in the body of the table. For example, the number 4,778 is the number of deaths other than battle deaths in the Marine Corps during World War II.

Ask the pupils to locate such data as:

- The total number in the Armed Forces during World War II
- The total casualties in the Marine Corps
- The battle deaths in the Navy
- Total deaths in all services
- The wounds not mortal in the Navy

Point out comparisons which may be made based on the data shown in the table:

- In which service did the number of deaths exceed the number of nonmortal wounds?
- The Army was approximately how many times as big as the Navy? the Marines? The Navy was how many times the size of the Marines? How many times as many nonmortal wounds as deaths were there in the Marine Corps?
- What fractional part of the total casualities in the Navy were not mortal? Deaths represented what part of the casualties in the Army?
- What percent of the men in each service were casualties? What percent of the total number in service were casualties?

After some comparisons have been made, show how generalizations can be drawn from the information in the table. Have the class discuss such generalizations as:

- The Marine Corp had the highest death rate and the highest total casualty rate: it was the most dangerous service to be in.
- The Navy had the lowest death rate and the lowest casualty rate: It was the safest service to be in. However, 63 percent of the casualties in the Navy were fatalities.

41. Interpreting Statistical Tables

Display a simple table such as that shown. Read the title, column heads and stub, and notice that the title does not tell that the information in the last column is included. Have a pupil read the note below the table to explain what the subscript means in the fifth column heading. Ask another pupil to explain what is indicated by the subscript "2" in the last column head, and why the letters "*n.a.*" are in the last space in that column.

For any year, read the number of buses, the number of other passenger motor vehicles, the number of trucks, and the total number of motor

vehicles. Comparing these numbers, the students should realize that the latter column is obtained by grouping the first three.

Find the increase from 1945 to 1952 for each type of vehicle.

Compute the percent of increase[1] from 1945 to 1952 for each type of vehicle. Which type of vehicle has increased by the greatest percent?

To determine whether the increase in surfaced roads kept pace with the increase in vehicles, divide the number of vehicles by the number of miles of highway. This ratio rose from 20 to 30 in the years 1945-51. Pupils should be cautioned that this number does not represent any normal load of traffic. It merely indicates that for every mile of surfaced highway, there was a certain number of vehicles in use that year. The variation of this ratio shows that the increase in roads did not keep pace with the increase in vehicles.

A check on this may be made by finding the percentages of increase in total vehicles and total miles of surfaced roads, and comparing them.

Motor Vehicles in Use 1945-1952

YEAR	BUSES	OTHER PASSENGER MOTOR VEHICLES	TRUCKS	TOTAL MOTOR VEHICLES[1]	MILES OF SURFACED HIGHWAY[2]
1945	81,875	25,691,434	4,834,742	30,608,051	1,495,000
1946	79,495	28,140,630	5,725,692	33,945,817	1,504,000
1947	89,000	30,880,861	6,521,369	37,022,230	1,544,000
1948	97,848	33,305,809	7,227,380	40,621,964	1,574,000
1949	97,855	36,329,819	7,692,569	44,120,243	1,617,000
1950	95,978	40,213,958	8,238,632	48,548,568	1,678,679
1951	94,689	42,573,818*a*	8,623,090	51,291,597	1,723,175
1952	92,201	43,908,472*a*	9,207,897	53,208,570	*n.a.*

Notes:

[1] Exclusive of motorcycles and publicly owned vehicles.

[2] Exclusive of city streets. Figures are estimated in part from available data.

a Includes sightseeing buses and school buses.

n.a. not available

[1] Percent of increase is discussed in activity 47.

Section VI

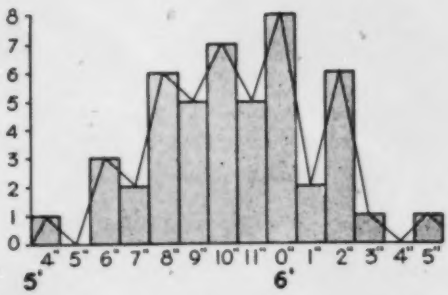
Statistical Terms

42. Discussion of Central Tendency

Usually data must be ordered in some logical manner before conclusions may be drawn regarding the data as a whole. For example, the weights of students within a class would be ordered from light to heavy, or heights would be ordered from short to tall. Once ordered, the data as a whole take on meaning; and the term "distribution" is used to refer to the relative arrangement of the data.

In the majority of cases, the data will tend to accumulate at a few of the positions in the order. In the case of the heights of high school boys, for example, the greatest numbers occur at heights near 5′ 10″, while heights of 6′ 4″ and 5′ 4″ are rare. A bar graph of boys' heights will, therefore, have the longest bars at heights which are close to 5′ 10″. If the tops of the bars were connected to form a line graph, an obvious bulge would be noted around 5′ 9″ to 6′ 0″. Thus, the shape of a distribution gives a "picture" of the data as a whole.

HEIGHT	NUMBER
5′ 4″	1
5′ 5″	0
5′ 6″	3
5′ 7″	2
5′ 8″	6
5′ 9″	5
5′ 10″	7
5′ 11″	5
6′ 0″	8
6′ 1″	2
6′ 2″	6
6′ 3″	1
6′ 4″	0
6′ 5″	1



Distribution of Heights of 47 Members of a High School Football Team

In the table the heights of the 47 members of a certain high school football team are indicated, and this information is plotted in the accompanying graph, which is to be displayed to the class. The distribution of the heights on one football team is not necessarily representative of the distribution obtained by consideration of all boys in high school. This distribution of player heights will be referred to in the following paragraphs as an example of a distribution of one specific set of data, not as an example of the distribution of the heights of all boys or of all high school football players.

43. Percentiles

Information regarding an ordered set of data is frequently given in terms of a "percentile." The relative position of a single member of the set can be expressed readily by means of the percentile. In fact, students may already be acquainted with the term because scores on standardized tests are often expressed in this form.

If John's score on a test is in the 72nd percentile it means that of all the papers written, 72 percent of them had scores equal to or less than that received by John. It also means that 28 percent received higher scores than John.

In order to acquaint the student with percentiles and their use, questions may be asked concerning the heights of the football players contained in activity 42. For example, a player in this group who is 6 feet tall ranks in what percentile? Answer: There are 47 players altogether and 37 of them are 6 feet tall or less, therefore one who is 6 feet tall would be in the 79th percentile $\left(\frac{37 \times 100}{47}\right)$. Additional questions of this type will further develop the percentile concept.

44. Illustrating the Meaning of Statistical Terms

To increase the statistical vocabulary of the pupils, give them copies of descriptions of various sets of data, such as the following, and discuss its correct interpretation.

In the House of Representatives of the U.S. Congress, the number of members from each State is apportioned according to the population of the State, and each State shall have at least one representative. In 1950, the following facts were true:

a. The number of representatives per State ranged from 1 to 43.
b. The average number of representatives was 9.
c. The median number was $7\frac{1}{2}$.
d. Quartile 1 was 2, and quartile 3 was 11.
e. The mode was 2.

The interpretation of statement *a* will be obvious to most pupils. Statement *b* can be further interpreted if we recall that the number of States was 48. This implies that the total number of representatives was 48 x 9, or 432. Actually there were 435, so that in this case the percent of error is less than 1 percent.

Statement *c* requires understanding of the word "median."

Statement *d* requires an extension of the concept of median: quartiles (Q) are points which divide an array of figures into quarters. Q_1 divides the lowest quarter from the upper three-fourths of the group. Q_2 divides the lower two-fourths from the upper two-fourths, and is therefore the same as the median. Q_3 divides the highest quarter from the lower three quarters. They are the 25th, 50th and 75th percentiles.

In this case, one fourth of the States had two representatives or less; one fourth had from two to seven; another fourth had from 8 to 11 and the last fourth had 11 or more.

Statement *e*, which uses the word "mode," indicates the value which is observed or which occurs the greatest number of times. Since the mode is given as 2, we know that there are more States having two representatives than having any other number of representatives.

45. Measuring Central Tendency

In general, the bulge of a distribution occurs near its middle; that is, the middle position in the order will usually have the value (or approximately the value) that occurs most frequently.

In the case of the heights of the 47 football players previously referred to, the 24th boy represents the central position when the players are ordered according to height. From the graph and table it is clear that the 24th boy is 5′ 10″ tall. If the boys were ordered in a line he would be the last member of the group who were 5′ 10″ tall and would be standing next to a boy 5′ 11″ tall. This one value at the middle position is a significant one and is called the "median" of the distribution.

Another value provides additional information about a distribution. The position in the order which has the greatest number of occurrences associated with it tends to lie in the center of the major bulge. The value at this position is called the "mode." Since eight football players have a height of 6 feet, then 6 feet is the mode of this particular distribution.

Both the median and mode tend to indicate where the major bulge, or "center," of the distribution occurs. The teacher should point out to the students that these values are just as important as the familiar "average,"* and that together these measures of "central tendency" provide considerable information about a distribution. The median, mode,

* Average, or arithmetic mean, is considered in activity 53.

and average are values which typify a set of data. In the case of the football players the average is approximately 5′ 10½″, the median is 5′ 10″ and the mode is 6′. These values indicate that the bulge in the distribution occurs around 5′ 10″—6′, and that it is virtually in the middle of the distribution.

The teacher should impress upon the class the dangers involved when only one of the measures of central tendency is used. Let's say, for example, that the members of a class received the following scores on an examination: 25, 27, 29, 29, 29, 30, 31, 97, 98, 100. The average of this set of data is 49.5, but this figure is not typical of the distribution. In this case the median and the mode are more representative of the distribution: the median is 29.5 and the mode is 29.

One more value will complete the picture. The "range" refers to the interval between the least and greatest members of a set. The range for the examination scores is 75 (100—25), while the range for the football players is 1′ 1″.

Whenever the median, mode and average lie near the middle of the range, the distribution will tend to be symmetric; that is, the quantities on either side of the median will generally match up. If the measures of central tendency are spread over the range, then the distribution will have several bulges or peaks, and none of the central measures is likely to be typical of the distribution.

The students should be made well aware of the dangers involved in accepting a single measure as representative of a distribution. It is sometimes said that statistics don't lie, but the truth of the matter is that statistics can lie and sometimes they are presented in such a way that they do lie If the average salary received in a small firm is said to be $6,000, but this figure is derived from two executive salaries over $12,000 coupled with six clerical salaries of $4,000 or less, then the "average" gives a false picture of the typical salary within the firm.

It is important that the students understand these measures of central tendency, because they are used to a great extent in everyday reading. In fact, the teacher may wish to have students search out mention of these terms in newspapers and weekly news magazines.

46. Understanding a Percentile Curve

Percentiles can be graphically represented by a curve sometimes referred to as a cumulative curve. Percentiles indicate the relative position of a score within a distribution by indicating the percent of scores which were equal to or less than a particular score. The percentile then is a cumulative quantity: the value for any one score depends upon the number of scores which have preceded it in the distribution. When

percentile values are plotted with respect to raw score values, a rising curve results.

In order to present the idea of a percentile curve to the class, a curve may be developed for the heights of the high school football players listed in activity 42. The table below contains a summary of the necessary information. For each height, expressed to the nearest inch, the number of players is indicated in the column headed "F" (frequency). The cumulative total is listed in the column labeled "Cum" (cumulative). This total is obtained for each height by adding the frequency for that height to the number of players with less height. The last column indicates the percentile of each height. This percent is obtained by dividing the cumulative total by the total number of players. These percentile figures are plotted in graph *a*.

As noted previously the percentile curve is a rising one. The example given is typical of the cumulative curve which may be expected from most small distributions. In cases where the data generally conform to a normal distribution[1] the percentile curve will be similar to illustration *b*.

A completed curve allows percentiles to be quickly approximated for a raw score or vice versa. Additional examples of percentile curves may be derived for distributions previously considered in class. Moreover, the manuals which accompany standardized tests usually contain percentile curves which assist in interpreting test results. Students may be asked to determine their own percentile ranks on a particular mathematics test, and a curve can be constructed on the board to consolidate the information for the entire class.

47. Interpreting Percent of Change

Present some statistics showing increases and decreases and percent change, such as the following sports records. Find the amount of change in each case; notice that it is compared only to the *former* amount and that the least *amount* of change is not necessarily the lowest percent of change.

Records of Some Olympic Games Champions (Men) 1896 vs 1952

EVENT	1896	1952	% CHANGE	AMOUNT OF CHANGE
Running High Jump	5′ 11¼″	6′ 8.32″	+13	
Pole Vault	10′ 9¾″	14′ 11.14″	+38	
100-Meter Run	12.0 sec.	10.4 sec.	—13	
110-Meter Hurdles	17 3/5 sec.	13.7 sec.	—22	

[1] Normal distribution, see activity 54.

Cumulative Totals and Percentiles of Football Players' Heights

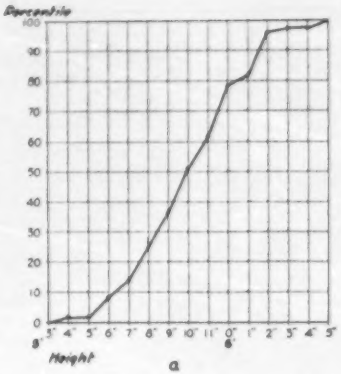
HEIGHT	F	CUM	$\frac{\text{CUM}}{\text{N}} = \%$
5′ 3″	0	0	0
5′ 4″	1	1	2
5′ 5″	0	1	2
5′ 6″	3	4	9
5′ 7″	2	6	13
5′ 8″	6	12	26
5′ 9″	5	17	36
5′ 10″	7	24	51
5′ 11″	5	29	62
6′ 0″	8	37	79
6′ 1″	2	39	83
6′ 2″	6	45	96
6′ 3″	1	46	98
6′ 4″	0	46	98
6′ 5″	1	47	100

Total Number 47

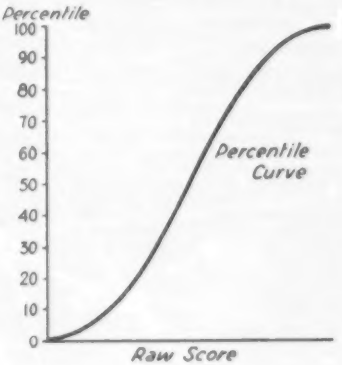
F = Frequency

CUM = Cumulative Totals

N = Total Number



a



b

Encourage students to interpret percents as fractions: + 13% means about ⅛ higher on the high jump; + 38% means about ⅜ higher on the pole vault. Likewise, — 13% means that the time for the 100-meter run was reduced by ⅛; — 22% means the time for 110-meter hurdles was reduced by more than one-fifth.

This may also be interpreted as the percent of the former record: for 100 meters, the record in 1952 was 87 percent of what it was in 1896. For the 110 meter hurdles, it was 78 percent of the 1896 record. The running high jump record is 113 percent, and the pole vault record is 138 percent of what the records were respectively in 1896.

48. Illustrating Decimal Parts of Percents of Change

Obtain from an almanac such figures as population changes in the States or Territories of the United States. Have the percent figures rounded off to the nearest whole percent and explained by pupils.

STATE OR TERRITORY	POPULATION IN 1950	POPULATION [1] IN 1956	INCREASE OR DECREASE	PERCENT OF INCREASE OVER 1950
Alaska	128,643	161,000		24.8
California	10,586,223	13,137,000		23.8
Hawaii	499,794	526,000		5.2
Massachusetts	4,690,514	4,766,000		1.6
New York	14,830,192	15,766,000		6.3
Puerto Rico	2,210,703	2,258,000		2.1
Vermont	377,747	369,000		— 2.4
Virgin Islands	26,665	30,000		10.0

[1] Census Bureau Estimate

If the percentages reported are checked in all these cases, slight discrepancies will be found. Discuss the possible explanations for this: the population figures of 1956 have been rounded; some percents may have been figured before rounding population figures.

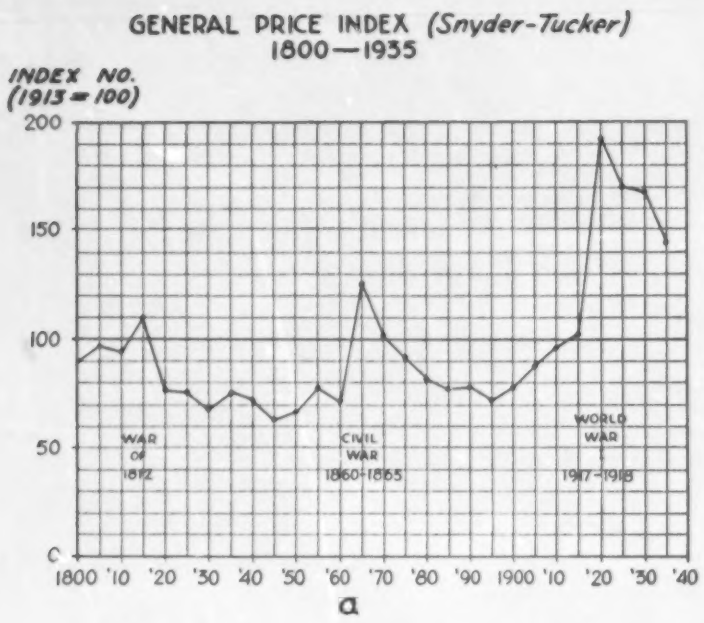
Display travel bureau pictures of each of these States and Territories. Consider which is growing the fastest, and why; which is growing slowest and why.

49. Interpreting Index Numbers

A time series of index numbers can be obtained from the *Statistical Abstract of the United States*, or some other comparable source. These numbers are best shown on a graph similar to that shown in *a*. The vertical scale is a percent scale, which always shows the 100 percent mark. The scale need not start at zero. Construct and display a chart similar to that shown.

The National Industrial Conference Board has published graphs in its "Roadmaps of Industry" series, which show index numbers; these graphs may be displayed also.

Discussion of index numbers should follow discussion of percent of increase or decrease. An index number is a magnitude expressed as a percent of the corresponding magnitude in some "base" period. The



a

magnitude during the base period is usually designated as 100. An item which had a magnitude 10 percent greater than the base would have an index number of 110. An item which had a magnitude 5 percent smaller than the base would have an index number of 95.

The average for some selected period of time is usually taken as the base for a given set of data.

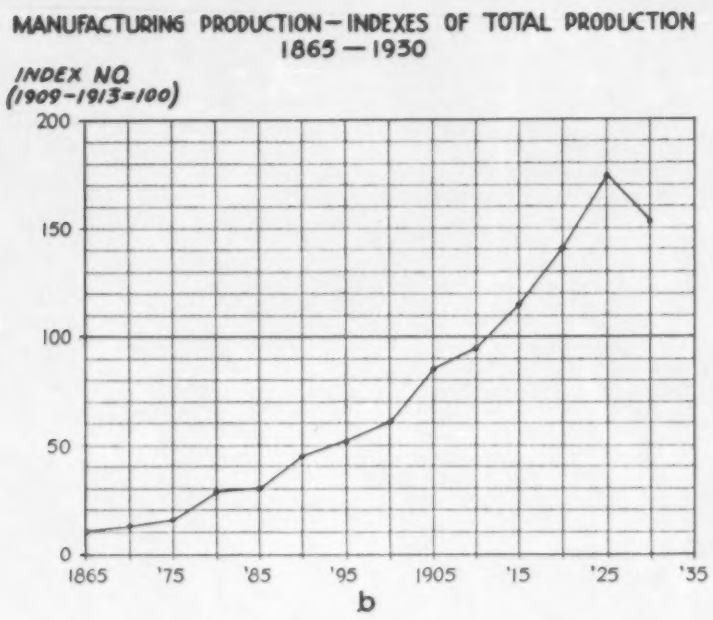
Ask pupils to read title, captions, scale and the data shown on the graph. Call attention to the way in which the base period is usually expressed on the graphs. In *a*, 1913 = 100 is to be interpreted, "the average for the year 1913 is designated as 100 on the graph and is used as the base."

Notice that extreme fluctuations can sometimes be explained by considering outside related factors.

Compare prices during any high period with prices during a low period.

Consider whether any trend is apparent; are there any cycles?

Display and consider graph *b*. Have the data explained and read. Notice the strong trend shown, and the fact that very few fluctuations are evident until the depression period.



50. Understanding Terminology

Parity. As an illustration of the use of the index concept, "parity" should be discussed. Parity is used primarily in relation to farming, and reference to parity is made almost daily in the newspapers.

The word "parity" means equality. It is used to indicate a type of index; that is, an index which represents an equal balance between income and expenditure. Farm prices are often expressed in terms of percent of parity prices. A parity price is one which will give a farmer on sale of his produce the same buying power as in a selected base year. If an item is sold at 70 percent of parity it would mean that the farmer could purchase only 70 percent of what he could have purchased as a result of a similar sale in the base year.

Farmers are not so much interested in how many dollars they get for wheat as in what they can buy with their wheat dollars. Even a price of $5 a bushel for wheat, if that price were 80 percent of parity would be less desirable to farmers than a price of $2 per bushel if the latter represented 100 percent of parity.

Ask students to clip newspaper articles which make reference to parity. Simple problems such as the following may be suggested:

If the purchasing power of the dollar decreases, will the parity price of corn increase or decrease?

Note: The base for parity used by the Federal Government has been the period 1910-14. Although this is still used for some commodities such as cotton, corn and wheat, a new base has been adopted for other items. This new base is obtained from the average price for the 10 years immediately preceding the year in question.

Sampling. The statistical term "sample" is used in much the same sense as one would normally use the word, as in the statement: "Sample the ice cream before serving it to the guests." In our normal speech "sample" means to taste or examine a small portion in order to determine the composition of the entire quantity. In the same way the statistician "samples" a large set of data to determine the characteristics of the entire set. This entire set is known as the "population." A sample is, then, a comparatively small representative portion of the "population."

Sampling techniques have become an important phase of statistics, because the vast amount of data with which researchers, government agencies and commercial concerns are confronted makes it virtually impossible for them to work with the basic populations. Students may be familiar with the public opinion polls which appear in the newspapers and news magazines. It may be noted to the class that the pollsters could not ask every person in the country to express his opinion about certain issues. Obviously only a very small portion of the people in the country could be requested to answer a group of questions, but this sample of the population must be representative of the entire Nation. This serves to introduce the idea of sampling, the need for workable numbers and the necessity of obtaining representative samples.

The teacher should warn students to be cautious in their acceptance of facts or conclusions based upon a sample. If the sample is not a truly representative one, then certainly the conclusions may be invalid. If a large number of graded examinations were stacked in a room so that the highest grades were on the top of the pile and the lowest grades at the bottom, then the use of the first 20 examinations picked off the top of the pile as a sample could result in many false conclusions.

Random Sample. A sequence is considered to be random if it follows no prescribed order. A random sample, therefore, is a small collection derived in a random manner from the basic population. The members of this type of sample are selected "blindly" in an effort to obtain a small representative set of data. Rather than ask all 800 pupils in a school to answer a confidential question in order to get an indication of sentiment on an issue, a random sample of at least 40 pupils could be expected to

yield a pattern of answers generally similar to that which would be obtained from the entire 800.

Students should be made aware of the care that must be taken in deriving a random sample. That is, the sample should be truly random. If a student were asked to place 30 numbers selected from 1 to 10 in a random order, the student would appreciate the difficulty in determining randomness. Any conscious attempt to place numbers in a random order is controlled by the individual's thoughts and as such they may follow a "prescribed order." Statisticians usually rely upon a mechanically derived random number table.

Random sampling is used in industry in order to maintain a close watch upon the quality of a product coming off an assembly line. Items chosen at random may be examined carefully by inspectors to assure that the finished product conforms to specifications.

In general, random samples are used when the population consists of members which are of the same general type. Thus the random sample is used to assure the control of quality of a certain product—every item is of the same general type.

Ask pupils to suggest other occasions when a random collection would be representative of a population.

Selected Sample. The random sample is not always the most representative which may be obtained from a population. Any given population is usually composed of several different subcategories. For example, the population of all major league baseball players may be broken into many different subcategories by considering such as teams, positions or performance averages. A sample of the major league population could be obtained by selecting a few members from each category. These selections may be made so that each category is represented in proportion to its size in the whole population. In such a selection some pitchers would be selected for sample, some infielders, some excellent hitters, some poor hitters, some average fielders etc.

This type of sample is usually employed by those who construct public opinion polls or surveys of a mixed population. Activity 51 suggests a selected sample from which the results of a class election may be predicted.

51. Engaging in Sampling Activities

Ask the class to suggest examples of occasions when sampling is desirable. For example, manufacturers before placing a new product on the market will often give samples of the product to a few people to determine their reaction. "Sneak" previews of movies are used for this purpose.

Students may be made more fully aware of the sampling process if they are asked to devise their own sample of a population. Suggestions may be sought for obtaining valid samples to determine weekly grocery expenditures by families in the United States or in the community, or to determine the amount and type of recreational reading done by local ninth graders.

52. Observing the Effects of Sampling

Two or three days before student elections for president of a student association, have pupils conduct a straw ballot among their friends or classmates. Each pupil may have a limited number of duplicated ballots to distribute and a sealed ballot box in which to collect the completed ballots.

Each pupil should be given a certain group to poll; these may be taken completely at random, or may be selected as representative or as non-representative groups.

Suggested representative cross sections in a certain school might be:

- 6 seniors, 7 juniors, 8 sophomores
- 8 college entrance students, 12 general, commercial or technical students
- 10 girls and 10 boys
- 3 students whose averages are A
  3 students whose averages are B
  18 students whose averages are C
  6 students whose averages are D
- 30 pupils in English classes representing all grades and groups. The proportions can be varied to conform to the proportions in each group.

Suggested groups which would not give a true cross section are:

- 10 members of Student Council
- 10 students whose averages are A
- 10 boys on school teams
- 10 girls in intramural athletics
- 10 pupils studying biology

Pupils should have a list or tally sheet for polling the representative groups, so that they may check off each classification as it is completed and there will be no duplication. During class, ballot boxes may be opened and votes from each group recorded. The votes from the "representative" groups may be totaled, and the percent of the total for each candidate computed. A prediction as to the outcome of the actual vote may be made from the percents.

The "nonrepresentative" groups may be considered in the same manner, and a forecast made concerning the final vote.

When the actual election is over, obtain the returns and compute the percent of votes actually received by each candidate. Perhaps the class estimate of representatives and nonrepresentative groups will need revision.

It should be pointed out that the individual samples of 10 persons are too small to serve as a basis for prediction and that at least 40 votes as a sample of the 800 pupils would be needed to make any kind of prediction. The accuracy of the prediction is related to the size of the sampling and also to the degree of homogeneity of the population sampled. A larger number of samples may be expected to yield greater accuracy.

53. Distinguishing between Arithmetic and Geometric Means

Considerable reference has been made to the average of a set of data. This is the term with which the student is most familiar and it is only right that this term be used in discussing measures of central tendency. However, once the idea of central tendency and its measures are well in mind it is important to call attention to other forms of "averages."

The statistician uses the term "arithmetic mean" to refer to what is commonly called the average. Although the arithmetic mean is by far the most common, other means are used in special cases. The fact that there are instances when the arithmetic mean is not typical of a set of data is an important one with which to acquaint students.

The harmonic mean and the geometric mean each have their uses in special distributions, but the geometric mean is the only one which should be considered. It is unlikely that the student will ever encounter the harmonic mean.

Suppose a sequence of terms is derived by multiplying the first term by a given number to get the second, and multiplying the second by the same quantity to find the third and so on. The short sequence 2, 4, 8 is an example of such a series of terms. Certainly a typical term for this sequence would be the number 4, but the arithmetic mean is approximately 4.7. The geometric mean, however, does give the value 4. If the three terms 2, 4 and 8 are multiplied together, the cube root of the product will be the geometric mean. Thus $2 \times 4 \times 8$ equals 64, and its cube root ($\sqrt[3]{64}$) is equal to 4.

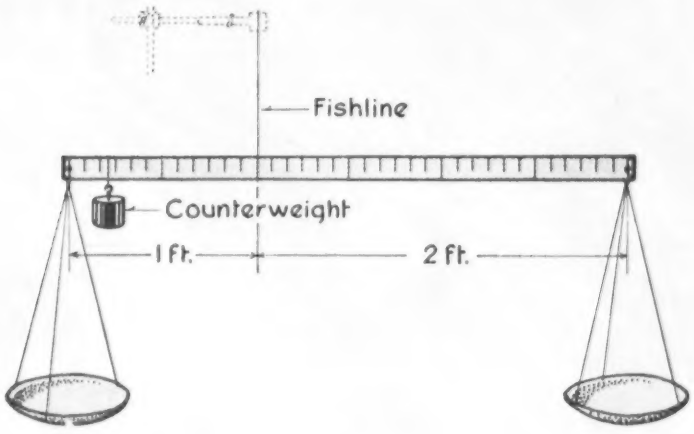
Similarly, if a sequence is formed by multiplying each term by 9 to get the succeeding term, the first two terms could be 1, 9. The arithmetic

mean of this two-termed sequence would be 5, but a more valid mean would be the geometric mean. The product of 1 and 9 is 9, and the square root of 9 gives the geometric mean 3. In each case the geometric mean is obtained by taking a root of the product of the terms. If the sequence consists of two terms, a square root is taken; if the sequence contains four terms, the fourth root is taken.

The use of a geometric mean in a short sequence can be illustrated by weighing an object with an off-center balance. A simple arrangement using a yardstick is shown in the diagram. The counterweight is used to balance the system before any weighing is done.

The object to be weighed may first be placed in the pan on the short arm and balanced with weights in the pan on the other arm. After this weighing, the arrangement is reversed: the object is placed in the pan on the longer arm and balanced with weights in the pan on the short arm. As a result two values are obtained, and a mean is used to eliminate any error caused by the difference in lengths of the arms.

If an object weighing 8 pounds were placed on the short arm of the scale in the figure it would be balanced by a weight of 4 pounds on the longer arm. With the arrangement reversed the object would be balanced by a weight of 16 pounds. It is obvious that the arithmetic mean would give a false weight, 10 pounds. The correct weight is given by the geometric mean: $4 \times 16 = 64$, $\sqrt{64} = 8$ pounds.



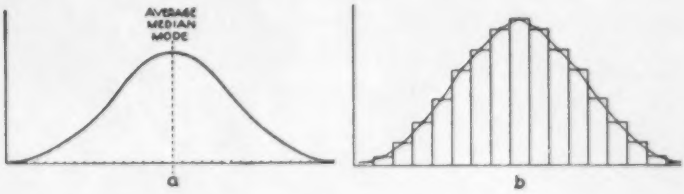
Section VII

Normal Distribution and Probability

54. Observing Normal Distribution

In discussing central tendency the teacher should stress the fact that a distribution may have many peaks or bulges and that the major concentration may possibly be located well away from the middle of the range. It is just as important, however, to acquaint students with distributions which may be expected to contain only one bulge, located in the middle of the range. A set of data which conforms to this pattern is said to be normally distributed.

In figure *a* a normal distribution is indicated by the bell-shaped curve. Note that the average, median and mode all occur at the center of the range and that the distribution is the same on either side of this center line. The curve which represents this entire distribution is called the normal curve.



It is perhaps unfortunate that a distribution of this type is referred to as being normal. The student may quickly assume that most distributions will take this shape and that those which do not must be considered abnormal. Such a conclusion is in error; indeed, the majority of distributions with which a student will come in contact will be quite different from the so-called "normal" shape. A normal distribution may be expected whenever a very large set of data is collected over a complete range. For example, if the heights of all the ninth grade boys in the State of New York were collected and indicated by a bar graph, and the mid points at the top of each bar connected by straight lines, a normal curve is approximated. (Figure *b*.)

It is important that examples show that normal distributions usually result from large sets of data which cover a complete range. Smaller distributions of the same types of data may be expected to approximate a normal distribution; the larger the set of data becomes the more the distribution may be expected to approximate a normal shape.

If approximately 100 observations are collected and plotted, a distribution which approximates the normal may be expected from areas such as the following: shoe sizes of boys or girls, ages of class members in months, student heights or student weights.

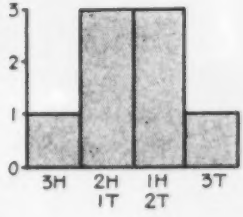
55. Developing the Probability Concept

Consideration of distribution and the normal curve leads easily into some basic concepts of probability. These concepts can be introduced very simply by discussing the flipping of a coin. It is obvious to most students that a coin has a 50-50 chance of falling heads once it has been flipped. The chance of a coin falling heads three times in succession is not quite so obvious and students find this an intriguing problem.

Students will find little difficulty in observing that one coin flipped three times in succession would be the same as flipping three coins all at once. The probability of flipping three consecutive heads would be identical to obtaining three heads from three coins flipped at the same time.

The coins should be labeled 1, 2 and 3. Then a table should be constructed such as that shown in figure *a*. A table is developed by considering the possible arrangements which can occur. If the first and second coins occur as heads, then the third can either be heads or tails; these account for the first two cases shown in the table. The first half of the

Possible Arrangements	Coin Nos.		
	#1	#2	#3
FIRST	H	H	H
SECOND	H	H	T
THIRD	H	T	H
FOURTH	H	T	T
FIFTH	T	H	H
SIXTH	T	H	T
SEVENTH	T	T	H
EIGHTH	T	T	T



3 HEADS	1
2 HEADS 1 TAIL	3
1 HEAD 2 TAILS	3
3 TAILS	1

a

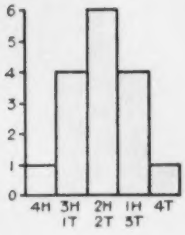
81

table is complete when all possible arrangements in which coin number 1 occurs as a head have been considered. The second half is complete when all arrangements involving a tail on the first coin have been considered. Once the table is finished the number of times each possible combination occurs is tabulated and the probability of three heads occurring is quickly seen.

There are eight possible arrangements of the coins but only one of these yields three heads. Therefore, there is a probability of $\frac{1}{8}$ (one out of eight) that three consecutive heads will occur. In other words, one can expect three heads to occur on one out of eight occasions. The probability of two heads and one tail occurring is $\frac{3}{8}$. It should be noted that the distribution of heads and tails is approximately normal, as may be seen in the bar graph. Figure *b* contains a table for four coins, as well as a bar graph showing the distribution.

The teacher may wish to pursue probability further by use of coin tables. A table for five coins may be constructed by the class, but this represents a practical limit because it involves 32 cases. Six coins would involve 64 cases.

Possible Arrangements	COIN NOS.			
	#1	#2	#3	#4
1st	H	H	H	H
2nd	H	H	H	T
3rd	H	H	T	H
4th	H	H	T	T
5th	H	T	H	H
6th	H	T	H	T
7th	H	T	T	H
8th	H	T	T	T
9th	T	H	H	H
10th	T	H	H	T
11th	T	H	T	H
12th	T	H	T	T
13th	T	T	H	H
14th	T	T	H	T
15th	T	T	T	H
16th	T	T	T	T



4 HEADS	1
3 HEADS 1 TAIL	4
2 HEADS 2 TAILS	6
1 HEAD 3 TAILS	4
4 TAILS	1

b

Additional problems may be examined from these probability tables. For example, if four coins are flipped, what is the chance of at least two of them being heads? From the table in *b* it is clear that one can expect four heads in $1/16$ of the times, three heads $4/16$ of the times, and two heads $6/16$ of the times. Thus at least two heads will occur $11/16$ of the times.

Still another type of problem may now be examined by the students. Consider the following: If a child's bank has in it two pennies, a nickel and a dime, what is the chance that the first three coins to be shaken out will yield exactly 12 cents? 16 cents? The solution to this problem may be obtained from table *c*, in which the possible arrangements of the first three coins are shown. There are four possible arrangements, but only one of them totals exactly 12 cents. The answer to the first part would then be $1/4$. Similarly the answer to the second part would be $2/4$ or $1/2$.

POSSIBILITIES	PENNY #1	PENNY #2	NICKLE	DIME	TOTAL AMOUNT
1	X	X	X		7¢
2	X	X		X	12¢
3	X		X	X	16¢
4		X	X	X	15¢

c

It is important to show students that probability is nothing more than the fraction of the total possibilities that a particular event may occur. Once this point has been made, the students may tackle most small problems of probability by means of tables.

Section VIII

Collecting and Preparing Statistical Data

Sources of Statistical Data

Offices of local, county or State governments publish reports and statistical tables which provide interesting data for local studies.

Results of polls conducted by such organizations as the American Institute of Public Opinion (Gallup poll) are reported in *The New York Times* and in certain other newspapers.

Libraries can supply some of the many statistical reports published by Federal agencies. Some of the reports which are most valuable as sources of current and historical data are:

The Decennial Census of the United States
The Monthly Labor Review ($6.25 per year)
Federal Reserve Bulletin ($6.00 per year)
Survey of Current Business ($4.00 per year)
Statistics of Income—1956
Part I—Individual ($.75)
Part II—Corporations ($1.25)
Handbook of Labor Statistics ($1.25)
Agricultural Statistics ($1.75—annual volume)

The above publications can be obtained from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C.

The following are important sources of data especially useful for mathematics teachers and may well be a part of high school libraries:

Statistical Abstract of the United States, 1954 (75th Anniversary Edition). For sale by Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. $3.75.

This issue contains over 1,000 pages of the latest statistics available on many subjects. Included are data on vital statistics, labor and employment, business, government, farming and prices. A 36-page bibliography of sources indicates where up-to-date figures are regularly obtainable. This volume is more useful than an almanac, because so many of the tables it includes are a time series.

Historical Statistics of the United States, 1789-1945. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. $2.50.

Data similar to those in the *Statistical Abstract* are included except that the series are extended backward in time in some cases to the 18th century. It contains excellent data for studying trends, cycles and fluctuations.

The Economic Almanac. A handbook of useful facts about business, labor and government in the United States and other areas. For sale by the National Industrial Conference Board, 460 Park Avenue, New York 22, N.Y. $4.

Information is similar to that in the statistical abstract, but is very complete in its industrial, manufacturing, business and labor statistics, as well as in data describing the economic status of the consumer. The almanac also has a glossary explaining many terms, such as "parity" and "index number."

•

56. Collecting Local Data

As classes collect and analyze data from the local community, this information should be recorded by the teacher and kept on file. After a few years, there will be a deposit of actual data available in the mathematics department. These data will serve as sources for interpretation and study. It is possible that the school newspaper would be interested in publishing the results.

Some possible subjects which might be areas for student research are listed below. Care should be taken not to ask for information in which wide variations occur.

- The heights of boys 14 years old, 15 years old, or the heights of senior girls or junior girls
- The weights of boys in the sophomore class
- The times of arising in the morning
- The times spent on buses each day
- The total allowances and wages received each week (estimate or average)
- Amounts of money spent on recreation each month by senior boys, senior girls, sophomore boys or 18-year-old girls
- The numbers of steps it takes to walk down the longest hall in the school building

Some data may be better obtained by the teacher than by the average pupil.

- The year's attendance for each student in the school or in one class. This can be obtained from the attendance registers in June.
- Marks on certain unit tests or final examinations
- Results of indirect measurements of inaccessible heights or distances, such as the height of a nearby structure
- The number of persons who buy a complete lunch in the cafeteria each day
- The number of books out on loan from the school library each day

Any other measure in which there will be a variation about some central tendency will lend itself to statistical study.

57. Appreciating the Qualities of a Good Questionnaire

Have each pupil make up three or four questions on topics in which he is interested. The form of the question and the space for the answer should be just as it would appear in the questionnaire. These may be collected, and any inappropriate or duplicate questions can be deleted by the teacher. A list of these questions may then be duplicated so that each pupil will have a copy for consideration.

To test the appropriateness of the questions, and the possibility of answering them briefly, have each student fill out the questionnaire. After the pupils have had an opportunity to react to the questions, discuss the whole questionnaire and the individual questions. Consider the following points:

- Is the questionnaire short enough so that a person is willing to finish all the questions?
- Does the question ask for information which is of interest and value?
- Is the question impersonal enough so that no one will hesitate to answer it?
- Can the question be answered accurately by the persons who will be given the questionnaire?
- Is the question clearly worded so that misunderstandings are avoided?
- Can the question be answered with one word or a very short phrase?
- Can the answers to a single question be classified easily, without too many categories?
- Is the format of the question such that answers can be tabulated easily?
- Is the questionnaire complete enough so that all data needed for analysis, such as the age or the grade of the person, will be obtained?

In order to answer the last question, a general plan for presenting the results should be considered, so that all necessary related data may be asked for.

In judging the questions on these points, it will be possible to eliminate many altogether, and to rewrite and change the format on others, until the questionnaire is in final acceptable form.

58. Using Questionnaires to Gather Raw Statistical Material

Frequently pupils in a class are not interested in the statistical material they are asked to graph. To help overcome this, have the class participate in preparing a questionnaire to be filled out by each pupil in his grade. List on the chalkboard the type of information indicated by discussion that they would like to compile about their class marks. If the list becomes too long, eliminate by show of hands those in which there is the least interest.

Then have class prepare questions as they would appear on a questionnaire and list these on the chalkboard. Consider each question, using the following criteria: Is it stated simply and concisely? Can it be misunderstood? Does it ask for exactly the information wanted? Can it be answered simply and directly? Are the method of answering and the place for the answer definite?

It may be necessary to add general directions in order to have the questionnaire filled out correctly. The teacher should watch to see that it does not become too long and also that it is not embarrassing to fill out. Names should not be asked for although an indication of sex may be necessary in using the data.

In order to secure a complete coverage of the entire population of the grade, the duplicated questionnaire could be filled out as an English class project or in homerooms. Graphs made from this information may be displayed so that all pupils participating can see the final results.

Typical questionnaire:

1. Indicate whether you are a boy or a girl. 1. _____
2. What is your height to the nearest inch? 2. _____
3. What is your weight? 3. _____
4. In what month is your birthday? 4. _____
5. How many brothers and sisters do you have? 5. _____
6. What is your favorite magazine? 6. _____
7. What is your favorite TV program? 7. _____
8. Who is your favorite recording artist or group? 8. _____
9. What is your favorite sport to play? 9. _____
10. What is your favorite sport to watch? 10. _____
11. What is your favorite hobby? 11. _____
12. What is your favorite season of the year? (fall, winter, spring, summer) 12. _____

59. Using Tally Sheets

To acquaint students with the techniques of gathering basic data, tally sheets may be made and class members assigned the task of gathering

Dessert	Number Selecting
Pudding	𝍸 𝍸
Apple	𝍸 𝍸 ////
Fruit cup	///
Pie	𝍸 𝍸 𝍸 𝍸 𝍸
Ice cream	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸

some raw data. A tally may be kept of the dessert selected by students if a choice is offered in the cafeteria lunch, as shown in the example. Book withdrawals from the school library may also be tallied.

Students should realize that although the tally sheet is a very simple device it permits raw data to be gathered in an orderly manner and it provides an immediate distribution. Tally sheets are used by traffic personnel in traffic checks. If a stoplight is being considered for a certain intersection, a traffic check is often conducted to determine the number of cars which pass through the intersection, the roads on which each car travels and the time of day it passes. The teacher may wish to conduct a simple corridor traffic check within the school as a class project.

60. Using Class Intervals

Very often the raw data which are to be gathered cover such a wide range that it would be impractical to consider the complete distribution. It is convenient to divide the range into intervals (called "class inter-

Class Interval	Tallies	Frequency
1 - 10	//	2
11 - 20	𝍸	5
21 - 30	////	4
31 - 40	𝍸 ///	8
41 - 50	𝍸 𝍸 ///	13
51 - 60	𝍸 𝍸 𝍸 𝍸	20
61 - 70	𝍸 𝍸 𝍸 𝍸 //	22
71 - 80	𝍸 𝍸 𝍸 𝍸 𝍸 /	26
81 - 90	𝍸 𝍸 𝍸 ////	19
91 -100	𝍸 𝍸 /	11

vals") and consider the distribution formed by the intervals. For example, if the scores on an examination were spread over a range from one to 100, it would probably suffice to consider the distribution formed by the intervals 1-10, 11-20, 21-30 and so on. The scores falling within these intervals would be tallied and the totals used to form the distribution to be considered. A typical distribution is shown in the illustration.

In order to show the need for class intervals other precise data covering a wide range should be considered by the pupils. The weights of students measured accurately to the nearest pound or half pound provide such data. Other examples could be derived from the number of pages in various books or the precise amounts on a number of paychecks.

Care should be exercised to assure that students realize the necessity of rounding off the data so that scores will fall within the correct intervals. Consideration of the weights of class members should offer opportunity to illustrate the necessity of rounding off.

The size of the class interval should be determined by the types of data to be considered, the range, the amount of data and the use to which the resulting distribution will be put.

Teachers may wish to devise sets of data for consideration by the class. Students should be asked to recommend a size for the class interval and then tally the data.

61. Selecting and Preparing the Means of Presentation

Data should be prepared for presentation in such a way that the most important properties and relationships will be effectively shown. The characteristics of the different types of graphs, tables, and statistical measures have been previously considered, but a short summary is included here.

Tables: Tables should be used to consolidate and classify many facts in a relatively small amount of space.

Statistical Measures: The range, median, mode and mean indicate the general form of a large distribution. Percentiles show the relative position of items within an ordered set of data.

Graphs: Graphs are used to add "eye appeal" to organized data and to provide illustrations of relationships.

- Line graph. Shows the change of one variable with respect to another.
- Bar graph. Compares sizes and quantities.
- Pictograph. Provides a variation of the bar graph.
- Circle and composite bar graphs. Illustrate component parts of a whole.

- Scatter diagrams. Show locations of a set of objects or the existence of a correlation between two sets of data.

Careful Preparation. Once data have been organized and examined and a means of presentation has been selected, care should be exercised to assure that the graph or table is accurately and neatly compiled. Toward this end, labels and headings should identify the data accurately in precise language. Footnotes and keys should be used to explain notations and permit readers to understand the material readily.

Reference can be made to the collection of tables and graphs which have been compiled in the class in order to illustrate the use of keys and explanatory notes. A roadmap also is a good example of information which is well keyed and explained.

Graphs should be of adequate size. Graphed data to be displayed to large audiences should be large enough to be seen clearly by all.

Color may be used to emphasize relationships shown on graphs. It may also be used to advantage to permit several sets of different data to be plotted on the same graph, in order to allow comparisons.

The student should be made aware of these small but important phases of graph preparation. The teacher may dwell upon these points whenever examples of good graphs are examined, and whenever students neglect to incorporate them into their own graphs.